**STATS 310A: Theory of Probability I** **Autumn 2016/17**

# Lecture 5: October 10

*Lecturer: Persi Diaconis* *Scribes: Kenneth Tay*

## 5.1 Law of the Iterated Logarithm

Recall the coin tossing set-up:

- $(\Omega, \mathcal{F}, P) = ([0,1], \text{Borel sets}, \text{Length})$.
- $r_i(\omega) := 2d_i(\omega) - 1$ (i.e. +1 or -1 "coin tosses").
- $S_n := \sum_{i=1}^{n} r_i$.

The Strong Law of Large Numbers (proved earlier in the course) says that, with probability 1,

$$\lim_{n\to\infty} \frac{S_n}{n} = 0.$$

This means that $S_n$ grows more slowly than $n$. What happens if we shrink the denominator? Can we approximate the growth of $S_n$ more exactly? The Law of the Iterated Logarithm gives us an answer to that question:

**Theorem 5.1 (Law of the Iterated Logarithm)** *Let $g(n) = \sqrt{2n\log\log n}$. Then, with probability 1,*

$$\overline{\lim} \frac{S_n}{g(n)} := \limsup \frac{S_n}{g(n)} = 1.$$

*In other words: for any $c > 1$, $\frac{S_n}{g(n)} < c$ for all large $n$, and for any $c < 1$, $\frac{S_n}{g(n)} > c$ infinitely often.*

**Note**: $\log\log 10^{100} \approx 5.2$, so $\log\log$ is essentially constant for anything practical.

In order to prove Theorem 5.1, we will need a few lemmas to help us. The first, known as the maximal inequality, will be proven below. The two lemmas following that were proven in Homework 2.

**Lemma 5.2 (Maximal Inequality)** *Let $M_n = \max(S_1, \ldots, S_n)$. Then for every integer $c \geq 1$,*

$$P\{M_n \geq c\} \leq P\{S_n \geq c\} + P\{S_n > c\} \leq 2P\{S_n \geq c\}.$$

**Note**: You can bound the maximum by the last sum!

**Proof:**

$$\begin{aligned} P\{M_n \geq c\} &= P\{M_n \geq c \text{ and } S_n \geq c\} + P\{M_n \geq c \text{ and } S_n < c\} \\ &=: I + II. \end{aligned}$$

Note that

$$I = P\{S_n \geq c\}.$$

For II: for each $1 \leq j \leq n$, let $F_j$ be the first time that $S_j = c$, i.e.

$$F_j = \{S_1 < c, S_2 < c, \ldots S_{j-1} < c, S_j = c\}.$$

Note that the $F_j$'s are disjoint. Then

$$\begin{aligned}
II &= P\{M_n \geq c \text{ and } S_n < c\} \\
&= \sum_j P\{F_j \text{ and } S_n - S_j < 0\} \\
&= \sum_j P\{F_j\}P\{S_n - S_j < 0\} && \text{(since } S_n - S_j \text{ only depends on coin flips after } j) \\
&= \sum_j P\{F_j\}P\{S_n - S_j > 0\} && \text{(by symmetry of coin flips)} \\
&= \sum_j^n P\{F_j \text{ and } S_n - S_j > 0\} \\
&= P\{S_n > c\}.
\end{aligned}$$

■

**Lemma 5.3** *For any* $\varepsilon > 0$,

$$P\left\{\frac{S_n}{n} \geq \varepsilon\right\} \leq 2\exp\left[-\frac{n\varepsilon^2}{2}\right].$$

**Lemma 5.4** *If* $\xi_n$ *is a sequence of real numbers such that* $\xi_n \to \infty$ *and* $\xi_n/\sqrt{n} \to 0$ *as* $n \to \infty$, *then*

$$P\{S_n \geq \sqrt{n}\xi_n\} \geq \exp\left[-\frac{(1+o(1))\xi^2}{2}\right].$$

We are now ready to prove the Law of Iterated Algorithm!

**Proof:**[ of Theorem 5.1] For each positive real number $c$, define the set

$$L_c = \{\omega : S_n(\omega) > cg(n) \text{ i.o.}\}.$$

Then the set of $\omega$ where the Law of the Iterated Algorithm holds can be explicitly written out as

$$L = \left(\bigcap_{c<1, c\in\mathbb{Q}} L_c\right) \cap \left(\bigcap_{c\geq 1, c\in\mathbb{Q}} L_c^c\right).$$

(Note that the $c$ are chosen such that $L$ is measurable.)

**Step 1: For $c > 1$, show that $P(L_c) = 0$.**

(Intuition: To use the 1st Borel-Cantelli Lemma, find some subsequence of events that have finite sum of probabilities.)

Let's say we have a subsequence $n_1 < n_2 < \dots$. (We will specify the $n_k$'s explicitly later.) Consider the event $\{\omega : S_n(\omega) > cg(n) \text{ for some } n \text{ s.t. } n_{k-1} < n \leq n_k\}$. Note that

$$\begin{aligned}\{\omega : S_n(\omega) > cg(n) \text{ for some } n \text{ s.t. } n_{k-1} < n \leq n_k\} \subseteq \{&M_{n_k} \geq cg(n_{k-1})\} \\ &=: A_k.\end{aligned}$$

However, by the Maximal Inequality, we have

$$\begin{aligned}
P\{M_{n_k} \geq cg(n_{k-1})\} &\leq 2P\{S_{n_k} \geq cg(n_{k-1})\} \\
&= 2P\left\{\frac{S_{n_k}}{n_k} \geq \frac{cg(n_{k-1})}{n_k}\right\} \\
&\leq 4\exp\left[-\frac{n_k}{2}\left(\frac{cg(n_{k-1})}{n_k}\right)^2\right] \qquad \text{(by Lemma 5.3)} \\
&= 4\exp\left[-\frac{c^2 g^2(n_{k-1})}{2n_k}\right] \\
&= 4\exp[-c^2 n_{k-1}\log\log n_{k-1}/n_k] \\
&= 4\left(\frac{1}{\log n_{k-1}}\right)^{c^2 n_{k-1}/n_k}.
\end{aligned}$$

Now we select $\{n_k\}$ carefully. Choose $\theta \in (1, c^2)$, and let $n_k = \lfloor \theta^k \rfloor$. Note that $n_k \nearrow \infty$. With this choice of $\{n_k\}$, we have

$$P\{M_{n_k} \geq cg(n_{k-1})\} \leq 4\left(\frac{1}{(k-1)\log\theta}\right)^{c^2/\theta(1+o(1))}$$

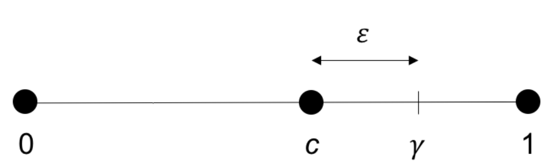
Since $c^2/\theta > 1$, we have $\sum P(A_k) < \infty$. Hence, we can use the 1st Borel-Cantelli Lemma to obtain $P(A_k \text{ i.o.}) = 0$.

Since $L_c = \{\omega : S_n(\omega) > cg(n) \text{ for some } n \text{ s.t. } n_{k-1} < n \leq n_k \text{ for infinitely many } k\} \subseteq A_k$ i.o., we have $P(L_c) = 0$, as required.

**Step 2: For $c < 1$, show that $P(L_c) = 1$.**

(Intuition: To use the 2nd Borel-Cantelli Lemma, find some subsequence of events which are independent and whose probabilities sum to infinity. We can't use the $\{S_n > cg(n)\}$ events directly as they are not independent.)

Choose a subsequence $1 < n_1 < \dots$ going to $\infty$. Let $\gamma = \frac{c+1}{2}$, $\varepsilon = \gamma - c$.



For any $k = 1, 2, \ldots$, define the sets

$$A_k = \left\{\omega : |S_{n_{k-1}}(\omega)| \leq \varepsilon g(n_k)\right\}, \qquad B_k = \left\{S_{n_k} - S_{n_{k-1}} \geq \gamma g(n_k)\right\}.$$

If $\omega \in A_k \cap B_k$, then $S_{n_k} = (S_{n_k} - S_{n_{k-1}}) + S_{n_{k-1}} \geq ((c+\varepsilon) - \varepsilon) g(n_k) = cg(n_k)$. Hence,

$$A_k \cap B_k \text{ i.o.} \subseteq \{S_{n_k} \geq cg(n_k) \text{ i.o.}\} \subseteq L_c.$$

We can go one step further to obtain

$$A_k \cap B_k \text{ i.o.} \supseteq \{A_k \text{ for all large } k\} \cap \{B_k \text{ i.o.}\}.$$

We claim that we can choose $\{n_k\}$ such that $P(\{A_k \text{ for all large } k\} \cap \{B_k \text{ i.o.}\}) = 1$, which implies that $P(L_c) = 1$. We prove this by showing $P\{A_k \text{ for all large } k\} = 1$ and $P\{B_k \text{ i.o.}\} = 1$.

From the bounds in Step 1, $|S_{n_{k-1}}| < 2g(n_{k-1})$ for all large $k$ with probability 1. For such an $\omega$,

$$\frac{S_{n_k}(\omega)}{g(n_k)} = \left[\frac{S_{n_k}(\omega)}{g(n_{k-1})}\right] \frac{g(n_{k-1})}{g(n_k)} \to 0$$

if we choose $\{n_k\}$ such that $n_k / n_{k-1}$ goes to infinity. With this, we have $P\{A_k \text{ for all large } k\} = 1$.

$$\begin{aligned}
P(B_k) &= P\{S_{n_k - n_{k-1}} \geq \gamma g(n_k)\} \\
&\geq \exp\left[-\frac{(1+o(1))\gamma^2 g^2(n_k)}{2(n_k - n_{k-1})}\right] \qquad \left(\text{Lemma 5.4 with } n = n_k - n_{k-1}, \xi_n = \frac{\gamma g(n_k)}{\sqrt{n_k - n_{k-1}}}\right) \\
&= \exp\left[-\frac{(1+o(1))\gamma^2 n_k \log\log n_k}{n_k - n_{k-1}}\right] \\
&= \left(\frac{1}{\log n_k}\right)^{(1+o(1))\gamma^2},
\end{aligned}$$

assuming $n_k / n_{k-1} \to \infty$.

Choose $\theta \in (1, 1/\gamma^2)$, let $n_k = \exp(k^\theta)$. $n_k$ goes to infinity, and $n_k/n_{k-1}$ goes to infinity. Thus,

$$\sum P(B_k) \geq \sum_k \frac{1}{k^{(1+o(1))\gamma^2 \theta}} = \infty.$$

Since the $B_k$'s are independent, we can apply the 2nd Borel-Cantelli Lemma to obtain $P\{B_k \text{ i.o.}\} = 1$.

**Step 3: Putting it all together.**

Recall that the set of $\omega$ where the Law of the Iterated Algorithm holds can be explicitly written out as

$$L = \left(\bigcap_{c<1, c\in\mathbb{Q}} L_c\right) \cap \left(\bigcap_{c\geq 1, c\in\mathbb{Q}} L_c^c\right).$$

From Step 1, every event in the second intersection has probability 1, and from Step 2, every event in the first intersection has probability 1. Since $L$ is a countable intersection of these sets, it must have probability 1 as well. ∎

Main ideas used in the proof:

- Bound $P\{M_n \geq c\} \leq 2P\{S_n \geq c\}$.
- Choose subsequences to make things independent.
- Large deviations bound: $P\{S_n > \xi\sqrt{n}\} \approx e^{-\xi_n^2/2}$.

**Note**: This theorem is due to Khintchine for coin tossing. Kolmogorov showed more generally that if $\{X_i\}_{i=1}^{\infty}$ iid with mean 0, finite variance $\sigma^2$, then

$$\limsup \frac{S_n}{\sqrt{2\sigma^2 n \log\log n}} = 1$$

with probability 1.