# STATS310A - Lecture 12

Persi Diaconis
Scribed by Michael Howes

10/28/21

## Contents



## 1 Kolmogorov's strong law of large numbers

Today we'll do some serious probability using many of our tools from measure theory.

**Theorem 1.** [Kolmogorov] *Let $X_1, X_2, \ldots$ be independent and identically distributed (i.i.d.) with finite mean $\mu = \mathbb{E}[X_1]$. Define*

$$S_n = \sum_{i=1}^{n} X_i,$$

*then $\frac{1}{n}S_n \to \mu$ almost surely.*

### 1.1 Proof

*Proof.* Write $X_i = X_i^+ - X_i^-$ and define $\mu^\pm = \mathbb{E}[X_1^\pm]$. Then $\mu = \mu^+ - \mu^-$ and so we can assume without loss of generality that $X_i \geq 0$.

Our next idea is to use truncation. Define

$$Y_i = X_i \delta_{\{X_i \leq i\}},$$

and let $T_n = \sum_{i=1}^n Y_i$. Our next idea is to look at subsequences. Fix $\alpha > 1$ and define $u_n = \lfloor \alpha^n \rfloor$. We now show that for all $\varepsilon > 0$,

$$\sum_{n=1}^{\infty} \mathbb{P}\left( \left| \frac{T_{u_n} - \mathbb{E}[T_{u_n}]}{u_n} \right| > \varepsilon \right) < \infty.$$

Note that by the i.i.d. assumption,

$$\begin{aligned}
\mathrm{Var}(T_{u_n}) &= \sum_{i=1}^{u_n} \mathrm{Var}(Y_i) \\
&\le \sum_{i=1}^{u_n} \mathbb{E}[Y_i^2] \\
&\le \sum_{i=1}^{u_n} \mathbb{E}[X_i^2 \delta_{\{X_i \le i\}}] \\
&= \sum_{i=1}^{u_n} \mathbb{E}[X_1^2 \delta_{\{X_1 \le i\}}] \\
&\le u_n \mathbb{E}[X_1^2 \delta_{\{X_1 \le u_n\}}].
\end{aligned}$$

By Chebyshev's inequality

$$\mathbb{P}\left(\left|\frac{T_{u_n} - \mathbb{E}[T_{u_n}]}{u_n}\right| > \varepsilon\right) \le \frac{\mathrm{Var}(T_{u_n})}{u_n^2 \varepsilon^2}.$$

Thus

$$\begin{aligned}
\sum_{n=1}^{\infty} \mathbb{P}\left(\left|\frac{T_{u_n} - \mathbb{E}[T_{u_n}]}{u_n}\right| > \varepsilon\right) &\le \sum_{n=1}^{\infty} \frac{\mathrm{Var}(T_{u_n})}{u_n^2 \varepsilon^2} \\
&\le \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{\mathbb{E}[X_1^2 \delta_{\{X_1 \le u_n\}}] u_n}{u_n^2} \\
&= \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \frac{\mathbb{E}[X_1^2 \delta_{\{X_1 \le u_n\}}]}{u_n} \\
&= \frac{1}{\varepsilon^2} \mathbb{E}\left[X_1^2 \sum_{n=1}^{\infty} \frac{1}{u_n} \delta_{\{X_1 \le u_n\}}\right].
\end{aligned}$$

We will now show that there exists $k$ such that if $x > 0$, then

$$\sum_{u_n \ge x} \frac{1}{u_n} \le \frac{k}{x}.$$

Let $n_x$ be the smallest integer such that $u_{n_x} \ge x$. One can then show that $\alpha^{n_x} \ge x$. It follows that

$$\begin{aligned}
\sum_{u_n \ge x} \frac{1}{u_n} &\le 2 \sum_{n \ge n_x} \alpha^{-n} \\
&= \frac{2\alpha^{-n_x}}{1-\alpha} \\
&= \frac{2}{1-\alpha} \times \frac{1}{\alpha^{n_x}} \\
&\le \frac{2}{1-\alpha} \times \frac{1}{x}.
\end{aligned}$$

Note that if $X_1 > 0$, then

$$\sum_{n=1}^{\infty} \frac{1}{u_n} \delta_{\{X_1 \le u_n\}} = \sum_{u_n \ge X_1} \frac{1}{u_n} \le \frac{k}{X_1}.$$

Since we have the convention $0 \cdot \infty = 0$, we can conclude that

$$\frac{1}{\varepsilon^2}\mathbb{E}\left[X_1^2 \sum_{n=1}^{\infty} \frac{1}{u_n} \delta_{\{X_1 \leq u_n\}}\right] \leq \frac{k}{\varepsilon^2}\mathbb{E}[X_1].$$

Thus Borel-Cantelli gives

$$\mathbb{P}\left(\left|\frac{T_{u_n} - \mathbb{E}[T_{u_n}]}{u_n}\right| > \varepsilon,\ i.o.\right) = 0.$$

If $(x_i)$ is any sequence of real numbers such that $x_i \to x$, then $\frac{1}{n}\sum_{i=1}^{n} x_i \to x$. By the monotone convergence theorem:

$$\mathbb{E}[Y_n] = \mathbb{E}[X_1 \delta_{\{X_1 \geq n\}}] \nearrow \mathbb{E}[X_1] = \mu.$$

Thus

$$\mathbb{E}\left[\frac{T_{u_n}}{u_n}\right] = \frac{1}{u_n}\sum_{i=1}^{u_n} \mathbb{E}[Y_i] \nearrow \mu.$$

Thus $\frac{T_{u_n}}{u_n} \to \mu$ with probability 1.

Thus we have proved our result on subsequences of the form $u_n = \lfloor \alpha^n \rfloor$ for truncated versions of $X_i$. We will first remove the truncation. Note that

$$\begin{aligned} \sum_{i=1}^{\infty} \mathbb{P}(X_i \neq Y_i) &= \sum_{i=1}^{\infty} \mathbb{P}(X_i \geq i) \\ &= \sum_{i=1}^{\infty} \mathbb{P}(X_1 \geq i) \\ &\leq \int_0^{\infty} \mathbb{P}(X_1 \geq t)dt \\ &= \mathbb{E}[X_1] \\ &< \infty. \end{aligned}$$

Thus Borel-Cantelli implies that $X_i = Y_i$ almost surely for sufficiently large $i$. Thus we can conclude $\frac{S_{u_n}}{u_n} \to \mu$ almost surely.

We will now use interpolation. We have $\frac{S_{u_n}}{u_n} \to \mu$ and we wish to conclude $\frac{S_k}{k} \to \mu$. For fixed $k$, let $n$ be such that $u_n \leq k \leq u_{n+1}$. Then

$$\frac{S_{u_n}}{u_n} \cdot \frac{u_n}{u_{n+1}} \leq \frac{S_k}{k} \leq \frac{S_{u_{n+1}}}{u_{n+1}} \cdot \frac{u_{n+1}}{u_n}.$$

We also have $\frac{u_n}{u_{n+1}} \to \frac{1}{\alpha}$ and $\frac{u_{n+1}}{u_n} \to \alpha$. Thus we have

$$\liminf_k \frac{S_k}{k} \geq \frac{1}{\alpha}\mu \ \text{ almost surely,}$$

and

$$\limsup_k \frac{S_k}{k} \leq \alpha\mu \ \text{ almost surely,}$$

for all $\alpha > 1$. Let $(\alpha_j)$ be a sequence such tht $\alpha_j \to 1$ and $\alpha_j > 1$. We then have

$$\frac{1}{\alpha_j}\limsup_k \frac{S_k}{k} \leq \mu \leq \alpha_j \liminf_k \frac{S_k}{k} \ \text{ almost surely.}$$

Thus $\liminf_k \frac{S_k}{k} = \limsup_k \frac{S_k}{k} = \mu$, almost surely. $\square$

### 1.2 The four T's

This was an example of using the four T's:

- Truncation,
- Tschebyscheff,
- Tsubsequences and
- InTerpolation.

All four of these are very useful ideas.

## 2 Remarks and stories

### 2.1 Remarks

(a) We showed that if $X_i$ are i.i.d., then

$$\frac{S_n}{n} \to \mu = \mathbb{E}[X_1] < \infty.$$

If $X_i \sim$ Cauchy i.e. $\mathbb{P}(X_n \leq x) = \int_{-\infty}^{x} \frac{1}{\pi(1+x^2)} dx$, then

$$\frac{X_1 + \ldots + X_n}{n} \sim \text{Cauchy}.$$

Thus we need $\mathbb{E}[X_1] < \infty$ for the strong law of large numbers.

(b) Independence:

i. Our proof works if we weaken the hypothesis from $(X_i)$ independent ot $(X_i)$ pairwise independent.

ii. We really only need that they are not too dependent. For example suppose $Y_i$ are i.i.d. with $\mathbb{E}[|Y_i|] < \infty$. If we define $X_1 = Y_1 + Y_2$, $X_2 = Y_2 + Y_3, \ldots$, then $(X_i)$ are not independent but they still satisfy a strong law of large numbers because a sum of $X_i$'s is essentially a sum of $Y_i$'s.

iii. Markov chains, martingales, stationary processes are all dependent and all have strong laws.

(c) Identically distributied? If none of the random variables are too wild, then the strong law is okay. For example we're safe if

$$\sum_{i=1}^{\infty} \frac{\text{Var}(X_i)}{i^2} < \infty.$$

We are not safe if $\text{Var}(S_n)$ is dominated by $\text{Var}(X_n)$. This might be the case if we are averaging stock prices or some other type of random variables which are sometimes very volatile.

### 2.2 Chebyshev's master thesis

In Chebyshev's master's thesis he proved that if $X_i \in \{0, 1\}$ are independent and $\mathbb{P}(X_i = 1) = p_i$, then $\mathbb{P}\left(\left|\frac{S_n}{n} - \frac{\sum_{i=1}^{n} p_i}{n}\right| > \varepsilon\right) \to 0$ and he had to prove this without Chebyshev's inequality (that came later).

## 2.3 Where's the beef?

We proved:

$$\boxed{\tfrac{S_n}{n} \to \mu \ a.s.}$$

which is a very nice statement but it doesn't tell us anything about finite $n$. What we really want is

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon \text{ for some } n \geq N\right) \leq f(N, \varepsilon),$$

for some function $f$. Such results are complicated and rare.