Notes 4 : Laws of large numbers

Math 733-734: Theory of Probability *Lecturer: Sebastien Roch*

References: [Fel71, Sections V.5, VII.7], [Dur10, Sections 2.2-2.4].

1 Easy laws

Let $X_1, X_2, \ldots$ be a sequence of RVs. Throughout we let $S_n = \sum_{k \leq n} X_k$.

We begin with a straighforward application of Chebyshev's inequality.

THM 4.1 (L^2 weak law of large numbers) *Let $X_1, X_2, \ldots$ be uncorrelated RVs, i.e., $\mathbb{E}[X_i X_j] = \mathbb{E}[X_i]\mathbb{E}[X_j]$ for $i \neq j$, with $\mathbb{E}[X_i] = \mu < +\infty$ and $\mathrm{Var}[X_i] \leq C < +\infty$. Then $n^{-1}S_n \to_{L^2} \mu$ and, as a result, $n^{-1}S_n \to_P \mu$.*

Proof: Note that

$$\begin{aligned}
\mathrm{Var}[S_n] &= \mathbb{E}[(S_n - \mathbb{E}[S_n])^2] = \mathbb{E}\left[\left(\sum_i (X_i - \mathbb{E}[X_i])\right)^2\right] \\
&= \sum_{i,j} \mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] = \sum_i \mathrm{Var}[X_i],
\end{aligned}$$

since, for $i \neq j$,

$$\mathbb{E}[(X_i - \mathbb{E}[X_i])(X_j - \mathbb{E}[X_j])] = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\mathbb{E}[X_j] = 0.$$

Hence

$$\mathrm{Var}[n^{-1}S_n] \leq n^{-2}(nC) \leq n^{-1}C \to 0,$$

that is, $n^{-1}S_n \to_{L^2} \mu$, and the convergence in probability follows from Chebyshev. ■

With a stronger assumption, we get an easy strong law.

THM 4.2 (Strong Law in L^4) *If the X_is are IID with $\mathbb{E}[X_i^4] < +\infty$ and $\mathbb{E}[X_i] = \mu$, then $n^{-1}S_n \to \mu$ a.s.*

Proof: Assume w.l.o.g. that $\mu = 0$. (Otherwise translate all X_is by μ.) Then

$$\mathbb{E}[S_n^4] = \mathbb{E}\left[\sum_{i,j,k,l} X_i X_j X_k X_l\right] = n\mathbb{E}[X_1^4] + 3n(n-1)(\mathbb{E}[X_1^2])^2 = O(n^2),$$

where we used that $\mathbb{E}[X_i^3 X_j] = 0$ by independence and the fact that $\mu = 0$. (Note that $\mathbb{E}[X_1^2] \leq 1 + \mathbb{E}[X_1^4]$.) Markov's inequality then implies that for all $\varepsilon > 0$

$$\mathbb{P}[|S_n| > n\varepsilon] \leq \frac{\mathbb{E}[S_n^4]}{n^4\varepsilon^4} = O(n^{-2}),$$

which is summable, and (BC1) concludes the proof.

■

The law of large numbers has interesting implications, for instance:

EX 4.3 (A high-dimensional cube is almost the boundary of a ball) *Let $X_1, X_2, \ldots$ be IID uniform on $(-1, 1)$. Let $Y_i = X_i^2$ and note that $\mathbb{E}[Y_i] = 1/3$, $\mathrm{Var}[Y_i] \leq \mathbb{E}[Y_i^2] \leq 1$, and $\mathbb{E}[Y_i^4] \leq 1 < +\infty$. Then*

$$\frac{X_1^2 + \cdots + X_n^2}{n} \to \frac{1}{3},$$

both in probability and almost surely. In particular, this implies for $\varepsilon > 0$

$$\mathbb{P}\left[(1-\varepsilon)\sqrt{\frac{n}{3}} < \|X^{(n)}\|_2 < (1+\varepsilon)\sqrt{\frac{n}{3}}\right] \to 1,$$

where $X^{(n)} = (X_1, \ldots, X_n)$. I.e., most of the cube is close to the boundary of a ball of radius $\sqrt{n/3}$.

2 Weak laws

In the case of IID sequences we get the following.

THM 4.4 (Weak law of large numbers) *Let $(X_n)_n$ be IID. A necessary and sufficient condition for the existence of constants $(\mu_n)_n$ such that*

$$\frac{S_n}{n} - \mu_n \to_P 0,$$

is

$$n\,\mathbb{P}[|X_1| > n] \to 0.$$

In that case, the choice

$$\mu_n = \mathbb{E}[X_1 \mathbb{1}_{|X_1| \leq n}],$$

works.

COR 4.5 (L^1 weak law) *If $(X_n)_n$ are IID with $\mathbb{E}|X_1| < +\infty$, then*

$$\frac{S_n}{n} \to_P \mathbb{E}[X_1].$$

Proof: From (DOM)

$$n\mathbb{P}[|X_1| > n] \leq \mathbb{E}[|X_1|\mathbb{1}_{|X_1|>n}] \to 0,$$

and

$$\mu_n = \mathbb{E}[X_1\mathbb{1}_{|X_1|\leq n}] \to \mathbb{E}[X_1].$$

■

Before proving the theorem, we give an example showing that the condition in Theorem 4.4 does not imply the existence of a first moment. We need the following important lemma which follows from Fubini's theorem. (Exercise.)

LEM 4.6 *If $Y \geq 0$ and $p > 0$, then*

$$E[Y^p] = \int_0^\infty p y^{p-1}\mathbb{P}[Y > y]\mathrm{d}y.$$

EX 4.7 *Let $X \geq e$ be such that, for some $\alpha \geq 0$,*

$$\mathbb{P}[X > x] = \frac{1}{x(\log x)^\alpha}, \quad \forall x \geq e.$$

(There is a jump at e. The choice of e makes it clear that the tail stays under 1.) Then

$$\mathbb{E}[X^2] = e^2 + \int_e^{+\infty} 2x\frac{1}{x(\log x)^\alpha}\mathrm{d}x \geq 2\int_e^{+\infty}\frac{1}{(\log x)^\alpha}\mathrm{d}x = +\infty, \quad \forall\alpha \geq 0.$$

(Indeed, it decays slower than $1/x$ which diverges.) So the L^2 weak law does not apply. On the other hand,

$$\mathbb{E}[X] = e + \int_e^{+\infty}\frac{1}{x(\log x)^\alpha}\mathrm{d}x = e + \int_1^{+\infty}\frac{1}{u^\alpha}\mathrm{d}u.$$

This is $+\infty$ if $0 \leq \alpha \leq 1$. But for $\alpha > 1$

$$\mathbb{E}[X] = e + \left.\frac{u^{-\alpha+1}}{-\alpha+1}\right|_1^{+\infty} = e + \frac{1}{\alpha - 1}.$$

Finally,

$$n\mathbb{P}[X > n] = \frac{1}{(\log n)^\alpha} \to 0, \quad \forall\alpha > 0.$$

(In particular, the WLLN does not apply for $\alpha = 0$.) Also, we can compute μ_n in Theorem 4.4. For $\alpha = 1$, note that (by the change of variables above)

$$\mu_n = \mathbb{E}[X \mathbb{1}_{X \leq n}] = e + \int_e^n \left(\frac{1}{x \log x} - \frac{1}{n \log n} \right) \mathrm{d}x \sim \log \log n.$$

Note, in particular, that μ_n may not have a limit.

2.1 Truncation

To prove sufficiency, we use truncation. In particular, we give a weak law for triangular arrays which does not require a second moment—a result of independent interest.

THM 4.8 (Weak law for triangular arrays) *For each n, let $(X_{n,k})_{k \leq n}$ be independent. Let b_n with $b_n \to +\infty$ and let $X'_{n,k} = X_{n,k} \mathbb{1}_{|X_{n,k}| \leq b_n}$. Suppose that*

1. $\sum_{k=1}^n \mathbb{P}[|X_{n,k}| > b_n] \to 0.$
2. $b_n^{-2} \sum_{k=1}^n \mathrm{Var}[X'_{n,k}] \to 0.$

If we let $S_n = \sum_{k=1}^n X_{n,k}$ and $a_n = \sum_{k=1}^n \mathbb{E}[X'_{n,k}]$ then

$$\frac{S_n - a_n}{b_n} \to_P 0.$$

Proof: Let $S'_n = \sum_{k=1}^n X'_{n,k}$. Clearly

$$\mathbb{P}\left[\left| \frac{S_n - a_n}{b_n} \right| > \varepsilon \right] \leq \mathbb{P}[S_n \neq S'_n] + \mathbb{P}\left[\left| \frac{S'_n - a_n}{b_n} \right| > \varepsilon \right].$$

For the first term, by a union bound

$$\mathbb{P}[S'_n \neq S_n] \leq \sum_{k=1}^n \mathbb{P}[|X_{n,k}| > b_n] \to 0.$$

For the second term, we use Chebyshev's inequality:

$$\mathbb{P}\left[\left| \frac{S'_n - a_n}{b_n} \right| > \varepsilon \right] \leq \frac{\mathrm{Var}[S'_n]}{\varepsilon^2 b_n^2} = \frac{1}{\varepsilon^2 b_n^2} \sum_{k=1}^n \mathrm{Var}[X'_{n,k}] \to 0.$$

■

Proof: (of sufficiency in Theorem 4.4) We apply Theorem 4.4 with $b_n = n$. Note that $a_n = n\mu_n$. Moreover,

$$\begin{aligned} n^{-1}\mathrm{Var}[X'_{n,1}] &\leq n^{-1}\mathbb{E}[(X'_{n,1})^2] \\ &= n^{-1}\int_0^\infty 2y\mathbb{P}[|X'_{n,1}| > y]\mathrm{d}y \\ &= n^{-1}\int_0^n 2y[\mathbb{P}[|X_{n,1}| > y] - \mathbb{P}[|X_{n,1}| > n]]\mathrm{d}y \\ &\leq 2\left(\frac{1}{n}\int_0^n y\mathbb{P}[|X_1| > y]\mathrm{d}y\right) \\ &\to 0, \end{aligned}$$

since we are "averaging" a function going to 0. Details in [D]. ■

The other direction is proved in the appendix.

3 Strong laws

Recall:

DEF 4.9 (Tail σ-algebra) *Let $X_1, X_2, \ldots$ be RVs on $(\Omega, \mathcal{F}, \mathbb{P})$. Define*

$$\mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \ldots), \quad \mathcal{T} = \bigcap_{n\geq 1} \mathcal{T}_n.$$

By a previous lemma, $\mathcal{T}$ is a σ-algebra. It is called the tail σ-algebra *of the sequence $(X_n)_n$.*

THM 4.10 (Kolmogorov's 0-1 law) *Let $(X_n)_n$ be a sequence of independent RVs with tail σ-algebra $\mathcal{T}$. Then $\mathcal{T}$ is $\mathbb{P}$-trivial, i.e., for all $A \in \mathcal{T}$ we have $\mathbb{P}[A] = 0$ or 1. In particular, if $Z \in \mathrm{m}\mathcal{T}$ then there is $z \in [-\infty, +\infty]$ such that*

$$\mathbb{P}[Z = z] = 1.$$

EX 4.11 *Let $X_1, X_2, \ldots$ be independent. Then*

$$\limsup_n n^{-1}S_n \quad \text{and} \quad \liminf_n n^{-1}S_n$$

are almost surely a constant.

3.1 Strong law of large numbers

THM 4.12 (Strong law of large numbers) *Let $X_1, X_2, \ldots$ be pairwise independent IID with $\mathbb{E}|X_1| < +\infty$. Let $S_n = \sum_{k\leq n} X_k$ and $\mu = \mathbb{E}[X_1]$. Then*

$$\frac{S_n}{n} \to \mu, \quad a.s.$$

If instead $\mathbb{E}|X_1| = +\infty$ then

$$\mathbb{P}\left[\lim_n \frac{S_n}{n} \text{ exists} \in (-\infty, +\infty)\right] = 0.$$

Proof: For the converse, assume $\mathbb{E}|X_1| = +\infty$. From Lemma 4.6

$$+\infty = \mathbb{E}|X_1| \leq \sum_{n=0}^{+\infty} \mathbb{P}[|X_1| > n].$$

By (BC2)

$$\mathbb{P}[|X_n| > n \text{ i.o.}] = 1.$$

Because

$$\frac{S_n}{n} - \frac{S_{n+1}}{n+1} = \frac{(n+1)S_n - nS_{n+1}}{n(n+1)} = \frac{S_n - nX_{n+1}}{n(n+1)} = \frac{S_n}{n(n+1)} - \frac{X_{n+1}}{n+1},$$

we get that

$$\{\lim_n n^{-1}S_n \text{ exists} \in (-\infty, +\infty)\} \cap \{|X_n| > n \text{ i.o.}\} = \emptyset$$

because, on that event, $S_n/n(n+1) \to 0$ so that

$$\left|\frac{S_n}{n} - \frac{S_{n+1}}{n+1}\right| > \frac{2}{3} \qquad \text{i.o.}$$

a contradiction. The result follows because $\mathbb{P}[|X_n| > n \text{ i.o.}] = 1$.

There are several steps in the proof of the $\Longrightarrow$ direction:

1. **Truncation.** Let $Y_k = X_k \mathbb{1}_{\{|X_k| \leq k\}}$ and $T_n = \sum_{k\leq n} Y_k$. (Note that the Y_i's are not identically distributed.) Since (by integrating and using Lemma 4.6)

$$\sum_{k=1}^{+\infty} \mathbb{P}[|X_k| > k] \leq \mathbb{E}|X_1| < +\infty,$$

(BC1) implies that it suffices to prove $n^{-1}T_n \to \mu$.

2. **Subsequence.** For $\alpha > 1$, let $k(n) = [\alpha^n]$. By Chebyshev's inequality, for $\varepsilon > 0$,

$$\begin{aligned}
\sum_{n=1}^{+\infty} \mathbb{P}[|T_{k(n)} - \mathbb{E}[T_{k(n)}]| > \varepsilon k(n)] &\leq \frac{1}{\varepsilon^2} \sum_{n=1}^{+\infty} \frac{\mathrm{Var}[T_{k(n)}]}{k(n)^2} \\
&= \frac{1}{\varepsilon^2} \sum_{n=1}^{+\infty} \frac{1}{k(n)^2} \sum_{i=1}^{k(n)} \mathrm{Var}[Y_i] \\
&= \frac{1}{\varepsilon^2} \sum_{i=1}^{+\infty} \mathrm{Var}[Y_i] \sum_{n:k(n)\geq i} \frac{1}{k(n)^2} \\
&\leq \frac{1}{\varepsilon^2} \sum_{i=1}^{+\infty} \mathrm{Var}[Y_i](Ci^{-2}) \\
&< +\infty,
\end{aligned}$$

where the next to last line follows from the sum of a geometric series and the last line follows from the next lemma—proved later:

LEM 4.13 *We have*

$$\sum_{i=1}^{+\infty} \frac{\mathrm{Var}[Y_i]}{i^2} \leq \mathbb{E}|X_1| < +\infty.$$

By (DOM) and (BC1), since ε is arbitrary, we have $\mathbb{E}[Y_k] \to \mu$ and

$$\frac{T_{k(n)}}{k(n)} \to \mu, \quad \text{a.s.}$$

3. **Sandwiching.** To use a sandwiching argument, we need a monotone sequence. Note that the assumption of the theorem applies to both X_1^+ and X_1^- and the result is linear so that we can assume w.l.o.g. that $X_1 \geq 0$. Then for $k(n) \leq m < k(n+1)$

$$\frac{T_{k(n)}}{k(n+1)} \leq \frac{T_m}{m} \leq \frac{T_{k(n+1)}}{k(n)},$$

and using $k(n+1)/k(n) \to \alpha$ we get

$$\frac{1}{\alpha}\mathbb{E}[X_1] \leq \liminf_m \frac{T_m}{m} \leq \limsup_m \frac{T_m}{m} \leq \alpha\mathbb{E}[X_1].$$

Since $\alpha > 1$ is arbitrary, we are done. But it remains to prove the lemma:
Proof: By Fubini's theorem

$$\begin{aligned}
\sum_{i=1}^{+\infty} \frac{\mathrm{Var}[Y_i]}{i^2} &\leq \sum_{i=1}^{+\infty} \frac{\mathbb{E}[Y_i^2]}{i^2} \\
&= \sum_{i=1}^{+\infty} \frac{1}{i^2} \int_0^\infty 2y\mathbb{P}[|Y_i| > y]\mathrm{d}y \\
&= \sum_{i=1}^{+\infty} \frac{1}{i^2} \int_0^\infty \mathbb{1}_{\{y \leq i\}} 2y\mathbb{P}[|Y_i| > y]\mathrm{d}y \\
&= \int_0^\infty \left(2y \sum_{i=1}^{+\infty} \frac{1}{i^2} \mathbb{1}_{\{y \leq i\}} \right) \mathbb{P}[|Y_i| > y]\mathrm{d}y \\
&\leq \int_0^\infty C' \mathbb{P}[|Y_i| > y]\mathrm{d}y \\
&\leq C' \mathbb{E}|X_1|,
\end{aligned}$$

where the second to last inequality follows by integrating. ■ ■

In the infinite case:

THM 4.14 (SLLN: Infinite mean case) *Let $X_1, X_2, \ldots$ be IID with $\mathbb{E}[X_1^+] = +\infty$ and $\mathbb{E}[X_1^-] < +\infty$. Then*

$$\frac{S_n}{n} \to +\infty, \quad \textit{a.s.}$$

Proof: Let $M > 0$ and $X_i^M = X_i \wedge M$. Since $\mathbb{E}|X_i^M| < +\infty$ the SLLN applies to $S_n^M = \sum_{i \leq n} X_i^M$. Then

$$\liminf_n \frac{S_n}{n} \geq \liminf_n \frac{S_n^M}{n} = \mathbb{E}[X_i^M] \uparrow +\infty,$$

as $M \to +\infty$ by (MON) applied to the positive part. ■

3.2 Applications

An important application of the SLLN:

THM 4.15 (Glivenko-Cantelli) *Let $(X_n)_n$ be IID and, for $x \in \mathbb{R}$,*

$$F_n(x) = \frac{1}{n} \sum_{k \leq n} \mathbb{1}\{X_k \leq x\},$$

be the empirical distribution function. *Then*

$$\sup_{x\in\mathbb{R}} |F_n(x) - F(x)| \to 0,$$

where F is the distribution function of X_1.

Proof: Pointwise convergence follows immediately from the SLLN. Uniform convergence then follows from the boundedness and monotonicity of F and F_n. See [D] for details. ■

References

[Dur10] Rick Durrett. *Probability: theory and examples*. Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge, 2010.

[Fel71] William Feller. *An introduction to probability theory and its applications. Vol. II.* Second edition. John Wiley & Sons Inc., New York, 1971.

A Symmetrization

To prove the other direction of the weak law, we use symmetrization.

DEF 4.16 *Let $X \sim F$. We say that X is* symmetric *if X and $-X$ have the same distribution function, that is, if at points of continuity we have $F(x) = 1 - F(-x)$ for all x.*

EX 4.17 (Symmetrization) *Let X_1 be a RV (not necessarily symmetric) and $\tilde{X}_1$, an independent copy. Then $X_1^\circ = X_1 - \tilde{X}_1$ is symmetric.*

LEM 4.18 *For all $t > 0$,*

$$\mathbb{P}[|X_1^\circ| > t] \leq 2\mathbb{P}[|X_1| > t/2]. \tag{1}$$

If m *is a median for X_1, i.e.,*

$$\mathbb{P}[X_1 \leq \mathrm{m}] \geq \frac{1}{2}, \quad \mathbb{P}[X_1 \geq \mathrm{m}] \geq \frac{1}{2},$$

and assume w.l.o.g. $\mathrm{m} \geq 0$ *then*

$$\mathbb{P}[|X_1^\circ| > t] \geq \frac{1}{2}\mathbb{P}[|X_1| > t + \mathrm{m}]. \tag{2}$$

Proof: For the first one, at least one of $|X_1| > t/2$ or $|\tilde{X}_1| > t/2$ must be satisfied. For the second one, the following are enough

$$\{X_1 > t + \mathrm{m}, \tilde{X}_1 \leq \mathrm{m}\} \cup \{X_1 < -t - \mathrm{m}, \tilde{X}_1 \geq -\mathrm{m}\},$$

and note that

$$\mathbb{P}[X_1 \geq -\mathrm{m}] \geq \mathbb{P}[X_1 \geq \mathrm{m}] \geq 1/2.$$

■

LEM 4.19 *Let $\{Y_k\}_{k\leq n}$ be independent and symmetric with $S_n = \sum_{k=1}^n Y_k$ and M_n equal to the first term among $\{Y_k\}_{k\leq n}$ with greatest absolute value. Then*

$$\mathbb{P}[|S_n| \geq t] \geq \frac{1}{2}\mathbb{P}\left[|M_n| \geq t\right]. \tag{3}$$

Moreover, if the Y_k's have a common distribution F then

$$\mathbb{P}[|S_n| \geq t] \geq \frac{1}{2}\left(1 - \exp(-n[1 - F(t) + F(-t)])\right). \tag{4}$$

Proof: We start with the second one. Note that

$$\mathbb{P}[|M_n| < t] \leq (F(t) - F(-t))^n \leq \exp\left(-n[1 - F(t) + F(-t)]\right).$$

Plug the latter into the the first statement.

For the first one, note that by symmetry we can drop the absolute values. Then

$$\mathbb{P}[S_n \geq t] = \mathbb{P}[M_n + (S_n - M_n) \geq t] \geq \mathbb{P}[M_n \geq t, (S_n - M_n) \leq 0]. \tag{5}$$

By symmetry, the four combinations $(\pm M_n, \pm(S_n - M_n))$ have the same distribution. Indeed M_n and $S_n - M_n$ are not independent but their sign is because M_n is defined by its absolute value and $S_n - M_n$ is the sum of the other variables. Hence,

$$\mathbb{P}[M_n \geq t] \leq \mathbb{P}[M_n \geq t, (S_n - M_n) \geq 0] + \mathbb{P}[M_n \geq t, (S_n - M_n) \leq 0],$$

and the two terms on the RHS are equal. Plugging this back into (5), we are done. ■

Going back to the proof of necessity:

Proof:(of necessity in Theorem 4.4) Assume that there is μ_n such that for all $\varepsilon > 0$

$$\mathbb{P}[|S_n - n\mu_n| \geq \varepsilon n] \to 0.$$

Note that

$$S_n^\circ = (S_n - n\mu_n)^\circ = \sum_{k\leq n} X_k^\circ.$$

Therefore, by (1), assuming w.l.o.g. $\mathrm{m} \geq 0$,

$$\begin{aligned}\mathbb{P}[|S_n - n\mu_n| \geq \varepsilon n] &\geq \frac{1}{2}\mathbb{P}[|S_n^\circ| \geq 2\varepsilon n] \\ &\geq \frac{1}{4}(1 - \exp(-n\mathbb{P}[|X_1^\circ| \geq 2n\varepsilon])) \\ &\geq \frac{1}{4}\left(1 - \exp\left(-\frac{1}{2}n\mathbb{P}[|X_1| \geq 2n\varepsilon + \mathrm{m}]\right)\right) \\ &\geq \frac{1}{4}\left(1 - \exp\left(-\frac{1}{2}n\mathbb{P}[|X_1| \geq n]\right)\right),\end{aligned}$$

for ε small enough and n large enough. Since the LHS goes to 0, we are done. ■

B St-Petersburg paradox

EX 4.20 (St-Petersburg paradox) *Consider an IID sequence with*

$$\mathbb{P}\left[X_1 = 2^j\right] = 2^{-j}, \quad \forall j \geq 1.$$

Clearly $\mathbb{E}[X_1] = +\infty$. *Note that*

$$\mathbb{P}[|X_1| \geq n] = \Theta\left(\frac{1}{n}\right),$$

(indeed it is a geometric series and the sum is dominated by the first term) and therefore we cannot apply the WLLN. Instead we apply the WLLN for triangular arrays to a properly normalized sum. We take $X_{n,k} = X_k$ *and* $b_n = n \log_2 n$. *We check the two conditions. First*

$$\sum_{k=1}^n \mathbb{P}[|X_{n,k}| > b_n] = \Theta\left(\frac{n}{n \log_2 n}\right) \to 0.$$

To check the second one, let $X'_{n,k} = X_{n,k}\mathbb{1}_{|X_{n,k}| \leq b_n}$ *and note*

$$\mathbb{E}[(X'_{n,k})^2] = \sum_{j=1}^{\log_2 n + \log_2 \log_2 n} 2^{2j}2^{-j} \leq 2 \cdot 2^{\log_2 n + \log_2 \log_2 n} = 2n \log_2 n.$$

So

$$\frac{1}{b_n^2}\sum_{k=1}^n \mathbb{E}[(X'_{n,k})^2] = \frac{2n^2 \log_2 n}{n^2(\log_2 n)^2} \to 0.$$

Finally,

$$a_n = \sum_{k=1}^{n} \mathbb{E}[X'_{n,k}] = n\mathbb{E}[X'_{n,1}] = n \sum_{j=1}^{\log_2 n + \log_2 \log_2 n} 2^j 2^{-j} = n(\log_2 n + \log_2 \log_2 n),$$

so that

$$\frac{S_n - a_n}{b_n} \to_P 0,$$

and

$$\frac{S_n}{n \log_2 n} \to_P 1.$$

THM 4.21 *Let $(X_n)_n$ be IID with $\mathbb{E}|X_1| = +\infty$ and $S_n = \sum_{k \leq n} X_k$. Let a_n be a sequence with a_n/n increasing. Then $\limsup_n |S_n|/a_n = 0$ or $+\infty$ according as $\sum_n \mathbb{P}[|X_1| \geq a_n] < +\infty$ or $= +\infty$.*

The proof uses random series and is presented in [D].

EX 4.22 (Continued) *Note that*

$$\mathbb{P}[|X_1| \geq n \log_2 n] = \Omega\left(\frac{1}{n \log_2 n}\right),$$

which is not summable. Therefore, by the previous theorem

$$\limsup_n \frac{S_n}{n \log_2 n} = +\infty, \quad \textit{a.s.}$$