Notes 5 : More on the a.s. convergence of sums

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

References: [Dur10, Sections 2.5]; [Wil91, Section 14.7], [Shi96, Section IV.4], [Dur10, Section 1.2].

1 Random series

1.1 Three-series theorem

We will give a second proof of the SLLN. This proof is based on random series and gives more precise information in some cases.

EX 5.1 Note that

\[ \sum_n \frac{1}{n} \text{ diverges,} \]

and

\[ \sum_n (-1)^{n+1} \frac{1}{n} = \sum_{m \geq 0} \left( \frac{1}{2m+1} - \frac{1}{2m+2} \right) = \sum_{m \geq 0} \frac{1}{2m(2m+1)} \text{ converges,} \]

in the sense that the partial sums converge. How about

\[ \sum_n \frac{Z_n}{n}, \]

where \( Z_n \in \{ \pm 1 \} \) are IID uniform? (Note that, with positive probability, you have long stretches of ones. On the other hand, you have +1s and −1s roughly half of the time.)

To answer the question, we prove:

THM 5.2 (Three-series theorem) Let \( X_1, X_2, \ldots \) be independent. Let \( A > 0 \) and \( Y_n = X_n 1_{\{|X_n| \leq A\}} \). In order for \( \sum_n X_n \) to converge a.s., it is necessary and sufficient that:

\[ \sum_{n=1}^{\infty} \mathbb{P}[|X_n| > A] < +\infty, \quad (1) \]


**Lecture 5: More on the a.s. convergence of sums**

\[
\sum_{n=1}^{\infty} \mathbb{E}[Y_n] \text{ converges,} \tag{2}
\]

and

\[
\sum_{n=1}^{\infty} \text{Var}[Y_n] < +\infty. \tag{3}
\]

**EX 5.3 (Continued)** Take \( A > 1 \). Then \( X_n = Y_n = n^{-1} Z_n \) and \( \mathbb{E}[Y_n] = 0 \) for all \( n \). Therefore it suffices to check (3). Note that

\[
\sum_n \text{Var}[X_n] = \sum_n \mathbb{E}[X_n^2] = \sum_n \frac{1}{n} < +\infty.
\]

1.2 Sufficiency

**Proof:** We claim it suffices to prove that

\[
\sum_n (Y_n - \mathbb{E}[Y_n]) \text{ converges.}
\]

Indeed, suppose that this is the case, then by (2) it follows that

\[
\sum_n Y_n \text{ converges.}
\]

By (1) and (BC1),

\[
\sum_n X_n \text{ converges.}
\]

So we prove:

**THM 5.4 (Zero-mean series theorem)** Let \( X_1, X_2, \ldots \) be independent with mean 0. If

\[
\sum_n \text{Var}[X_n] < +\infty
\]

then \( S_n = \sum_{k \leq n} X_k \) converges a.s.

**Proof:** For \( M > 0 \), let

\[
w_M \equiv \sup_{n, m \geq M} |S_n - S_m| \downarrow \omega_{\infty}.
\]

By a Cauchy-type argument it suffices to prove that, for all \( \varepsilon > 0 \),

\[
\mathbb{P}[w_{\infty} > 2\varepsilon] = \mathbb{P}[\cap_M \{ w_M > 2\varepsilon \}] = \lim_{M} \mathbb{P}[w_M > 2\varepsilon] = 0,
\]

by continuity.

To argue about the \( \sup \) we prove the following:
Lecture 5: More on the a.s. convergence of sums

**THM 5.5 (Kolmogorov’s maximal inequality)**  
Let $X_1, X_2, \ldots$ be independent with 0 mean and finite variance. Then

$$
P \left[ \max_{1 \leq k \leq n} |S_k| \geq x \right] \leq \frac{\text{Var}[S_n]}{x^2}.
$$

Note that this is a strengthening of Chebyshev’s inequality for sums of independent RVs.

**Proof:** The trick is to divide the event of interest into when it first occurs. Let

$$A_k = \{ |S_k| \geq x \text{ but } |S_j| < x \text{ for } j < k \}.$$

Then, using independence of $S_k 1_{A_k}$ and $(S_n - S_k)$ and the fact that $E[S_n - S_k] = 0$,

$$
E[S_n^2] \geq \sum_{k=1}^{n} E[S_n^2; A_k] \\
= \sum_{k=1}^{n} \left( E[S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2; A_k] \right) \\
\geq \sum_{k=1}^{n} \underbrace{E[S_k^2; A_k]}_{\geq x^2P[A_k]} \\
\geq \sum_{k=1}^{n} x^2 P[A_k] \\
= x^2 P \left[ \max_{1 \leq k \leq n} |S_k| \geq x \right].
$$

Going back to the previous proof, note that Kolmogorov’s maximal inequality implies

$$
P \left[ \max_{M \leq m \leq N} |S_m - S_M| > \varepsilon \right] \leq \frac{\text{Var}[S_N - S_M]}{\varepsilon^2} \leq \frac{\sum_{m \geq M+1} \text{Var}[X_m]}{\varepsilon^2},
$$

so taking a limit when $N \to +\infty$ then $M \to +\infty$ and using continuity

$$
P \left[ \sup_{m \geq M} |S_m - S_M| > \varepsilon \right] \to 0, \quad \text{as } M \to +\infty.
$$

Finally, note that

$$
P[w_M > 2\varepsilon] \leq P \left[ \sup_{m \geq M} |S_m - S_M| > \varepsilon \right] \to 0,
$$

using $|S_n - S_m| \leq |S_n - S_M| + |S_m - S_M|$.
1.3 Necessity

Proof: Suppose \( \sum_n X_n \) converges. Then \( X_n \to 0 \) and by (BC2), for all \( A > 0 \), (1) holds and \( \sum_n Y_n \) converges as well. We use symmetrization. If \( Z \) is a RV and \( \tilde{Z} \) is an independent copy. Then we let \( Z^O = Z - \tilde{Z} \). We will show below that the convergence of a zero-mean, uniformly bounded random series implies the convergence of the second moments. Then

\[
\sum_n Y_n^O \text{ converges} \implies \sum_n \text{Var}[Y_n] = \frac{1}{2} \sum_n \text{Var}[Y_n^O] < +\infty
\]

\[
\implies \sum_n (Y_n - E[Y_n]) \text{ converges}
\]

\[
\implies \sum_n E[Y_n] \text{ converges.}
\]

The statement above will follow from the following maximal inequality:

**THM 5.6** Let \( X_1, X_2, \ldots \) be independent with zero mean, finite variance and \( P[X_i \leq c] = 1 \) for some \( c > 0 \). Then

\[
P\left[ \max_{1 \leq k \leq n} |S_k| \geq x \right] \geq 1 - \frac{(c + x)^2}{\text{Var}[S_n]}.
\]

Proof: Let \( A \) be the event above and define \( A_k \) as before. On the one hand, noting that on \( A_k \) we have \( |S_k| \leq c + x \) and arguing as before,

\[
E[S_n^2; A] = \sum_k E[S_k^2 + (S_n - S_k^2); A_k] \leq P[A] \left( (c + x)^2 + \sum_{k=1}^n E[X_k^2] \right) = P[A] \left( (c + x)^2 + E[S_n^2] \right).
\]

On the other hand,

\[
E[S_n^2; A] = E[S_n^2] - E[S_n^2; A^c] \geq E[S_n^2] - x^2 P[A^c].
\]

Rearranging gives the result. Finally if we had \( \sum_n \text{Var}[Y_n] = +\infty \), the previous inequality (used as before) would give

\[
P\left[ \sup_{k \geq M} |S_n - S_M| > \varepsilon \right] = 1,
\]

for all \( M > 0 \)—a contradiction.
1.4 Applications

1.4.1 A second proof of the SLLN

To see the connection between the convergence of series and the law of large numbers, we begin with the following lemma (whose proof is in the appendix).

LEM 5.7 (Kronecker’s lemma) If $a_n \uparrow +\infty$ and $\sum_n x_n/a_n$ converges then

$$\frac{1}{a_n} \sum_{m=1}^n x_m \to 0.$$ 

THM 5.8 (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

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

Proof: We already proved that it suffices to show $n^{-1}T_n \to \mu$ where $Y_k = X_k1_{\{|X_k| \leq k\}}$ and $T_n = \sum_{k \leq n} Y_k$, that $\mathbb{E}[Y_k] \to \mu$, and that

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

That implies, from Theorem 5.4 (or from the Three-Series Theorem with $A > 2$),

$$\sum_k \frac{Y_k - \mathbb{E}[Y_k]}{k} \text{ converges a.s.}$$

1.4.2 Rates of convergence

Under a stronger assumption, we have the following:

THM 5.9 (Rate of convergence) Let $X_1, X_2, \ldots$ be IID with $\mathbb{E}|X_1| = \mu < +\infty$ and $\text{Var}[X_1] = \sigma^2 < +\infty$. Then $S_n = \sum_{k \leq n} X_k$ satisfies

$$\frac{S_n - n\mu}{n^{1/2}(\log n)^{1/2+\varepsilon}} = \frac{n^{1/2}}{(\log n)^{1/2+\varepsilon}} \left( \frac{S_n}{n} - \mu \right) \to 0, \quad \text{a.s.}$$
Lecture 5: More on the a.s. convergence of sums

Proof: Take $\mu = 0$ w.l.o.g. Define $a_n = n^{1/2} (\log n)^{1/2+\varepsilon}$. Then

$$\sum_{n \geq 2} \text{Var} \left[ \frac{X_n}{a_n} \right] = \sum_{n \geq 2} \frac{\sigma^2}{n (\log n)^{1+2\varepsilon}}.$$  

Recalling that

$$\int \frac{1}{x (\log x)^\alpha} dx = (1 - \alpha) \frac{1}{(\log x)^{\alpha-1}} + C,$$

for $\alpha > 1$ we get that the previous series is finite. Hence, from Theorem 5.4, $\sum_n X_n / a_n$ converges and $a_n^{-1} \sum_{k \leq n} X_k$ goes to 0.

\[\square\]

2 Law of the iterated logarithm

Let $(X_n)_n$ be IID with $\mathbb{E}[X_1] = 0$ and $\text{Var}[X_1] = \sigma^2 < +\infty$ and let $S_n = \sum_{k \leq n} X_k$. Using random series results, we proved that

$$\frac{S_n}{\sqrt{n \log^{1/2+\varepsilon} n}} \to 0, \text{ a.s.}$$

(On the other hand, we will prove later that the distribution of $n^{-1/2} S_n$ converges to a non-trivial limit.)

In fact:

**THM 5.10 (Law of Iterated Logarithm (LIL))** Let $(X_n)_n$ be IID with $\mathbb{E}[X_1] = 0$ and $\text{Var}[X_1] = \sigma^2 < +\infty$ and let $S_n = \sum_{k \leq n} X_k$. Then

$$\limsup_n \frac{S_n}{\sqrt{2\sigma^2 n \log \log n}} = 1, \text{ a.s.}$$

We prove the result in a special case: assume from now on that $X_1 \sim N(0, 1)$. (The more general case uses Brownian motion and is described in [Dur10, Section 8.8].) But first:

2.1 Reminder: Normal distribution

Recall:

**DEF 5.11 (Density)** A RV $X$ has a probability density function (PDF) $f_X : \mathbb{R} \to [0, +\infty)$ if

$$\forall B \in \mathcal{B}(\mathbb{R}), \mathbb{P}[X \in B] = \int_B f_X(y) dy.$$
Then, by Exercise 1.6.8, if $\mathbb{E}|h(X)| < +\infty$ then

$$\mathbb{E}[h(X)] = \int h(x)f_X(x)dx.$$ 

Then:

**DEF 5.12 (Gaussian)** A RV $X$ has a normal distribution with mean $\mu$ and variance $\sigma^2$ (denoted $X \sim N(\mu, \sigma^2)$) if it has density

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right).$$

Note in particular that

$$\left(\int_{-\infty}^{+\infty} f_X(x)dx\right)^2 = \int_{-\infty}^{+\infty} \frac{1}{2\pi\sigma^2} \exp\left(-\frac{(x - \mu)^2}{2\sigma^2}\right) dx dy$$

$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \exp\left(-\frac{x^2 + y^2}{2}\right) d\tilde{x} d\tilde{y}$$

$$= \frac{1}{2\pi} \int_{0}^{+\infty} \exp\left(-\frac{r^2}{2}\right) r dr$$

$$= \frac{1}{2\pi} \int_{0}^{+\infty} \exp(-u) du$$

$$= 1,$$

where we used polar coordinates on the third line. Also,

$$\mathbb{E}[X - \mu] = \int_{-\infty}^{+\infty} (x - \mu)f_X(x)dx$$

$$= \sigma \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) dz$$

$$= 0,$$

by symmetry and, by linearity of expectation and using polar coordinates again,

$$\text{Var}[X] = \mathbb{E}[(X - \mu)^2]$$

$$= \frac{\sigma^2}{2} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left(x^2 + y^2\right) \frac{1}{2\pi} \exp\left(-\frac{x^2 + y^2}{2}\right) d\tilde{x} d\tilde{y}$$

$$= \frac{\sigma^2}{2} \int_{0}^{+\infty} r^2 \exp\left(-\frac{r^2}{2}\right) r dr$$

$$= \sigma^2\int_{0}^{+\infty} u \exp(-u) du,$$

$$= \sigma^2.$$
by integration by parts. That is, \( \mathbb{E}[X] = \mu \) and \( \text{Var}[X] = \sigma^2 \).

Let \( X \sim N(\mu_1, \sigma_1^2) \) and \( Y \sim N(\mu_2, \sigma_2^2) \) be independent normal RVs. By our convolution formula, the PDF of \( X + Y \) is

\[
 f_{X+Y}(z) = \int_{-\infty}^{+\infty} f_X(z-y) f_Y(y) \, dy \\
 = \int_{-\infty}^{+\infty} \frac{1}{2\pi \sigma_1 \sigma_2} \exp \left( -\frac{(z-y-\mu_1)^2}{2\sigma_1^2} + \frac{(y-\mu_2)^2}{2\sigma_2^2} \right) \, dy \\
 = \cdots = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp \left( -\frac{(z-\mu_1-\mu_2)^2}{2(\sigma_1^2 + \sigma_2^2)} \right),
\]

that is, \( X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2) \). (See [Dur10, Example 2.1.4] for the computations.)

2.2 Proof of LIL in Gaussian case

We start with two lemmas of independent interest. First, because we are dealing with a \( \lim sup \) we will need a maximal inequality.

LEM 5.13 Let \((X_n)_n\) be independent and symmetric (that is, \(X_1\) and \(X_1\) have the same distribution). Then for all \(a \in \mathbb{R}\)

\[
 \mathbb{P} \left[ \max_{1 \leq k \leq n} S_n > a \right] \leq 2 \mathbb{P}[S_n > a].
\]

This is similar to Kolmogorov’s maximal inequality. However, we will get a stronger bound in the Gaussian case below. First, we prove the lemma.

**Proof:** Let \( A \) be the set in bracket on the LHS and \( B \) be the set in bracket on the RHS. Define

\[
 A_k = \{S_k > a \text{ but } S_j \leq a, \forall j < k\}.
\]

Note that

\[
 \mathbb{P}[B] \geq \sum_{k \leq n} \mathbb{P}[A_k \cap B]. \tag{4}
\]

Moreover,

\[
 \mathbb{P}[A_k \cap B] \geq \mathbb{P}[A_k \cap \{S_n \geq S_k\}] \\
 = \mathbb{P}[A_k] \mathbb{P}[S_n - S_k \geq 0] \\
 \geq \frac{1}{2} \mathbb{P}[A_k],
\]

by symmetry. Plugging back into (4) gives the result. \( \blacksquare \)
LEM 5.14 For \( x > 0 \),
\[
(x^{-1} - x^{-3})e^{-x^2/2} \leq \int_x^{+\infty} e^{-y^2/2} dy \leq x^{-1}e^{-x^2/2}.
\]

Proof: By the change of variable \( y = x + z \) and using \( e^{-z^2/2} \leq 1 \)
\[
\int_x^{+\infty} e^{-y^2/2} dy \leq e^{-x^2/2} \int_0^{+\infty} e^{-xz} dz = e^{-x^2/2}x^{-1}.
\]
For the other direction, by differentiation
\[
\int_x^{+\infty} (1 - 3y^{-4})e^{-y^2/2} dy = (x^{-1} - x^{-3})e^{-x^2/2}.
\]

We come back to the proof of the main theorem.

Proof: Define
\[
h(n) = \sqrt{2n \log \log n}.
\]

Upper bound. The upper bound follows from (BC1). Let \( K > 1 \) (close to 1) and \( c_n = Kh(K^{n-1}) \). In words we want to show that \( S_k/h(k) \) is smaller than \( K \) eventually. By the lemmas above
\[
\Pr \left[ \max_{k \leq K^n} S_k > c_n \right] \leq 2 \Pr [S_{K^n} > c_n]
=
2 \Pr \left[ \frac{S_{K^n}}{K^{n/2}} > \frac{c_n}{K^{n/2}} \right]
\leq 2 \frac{1}{\sqrt{2\pi}} \frac{K^{n/2}}{c_n} \exp \left( -\frac{c_n^2}{2K^n} \right).
\]
Note that
\[
\frac{c_n^2}{2K^n} = K \log \log K^{n-1} = K[\log(n-1) + \log \log K] = \log(n-1)^K + \log(\log K)^K.
\]

For \( n \) large enough
\[
\Pr \left[ \max_{k \leq K^n} S_n > c_n \right] \leq \exp \left( -\frac{c_n^2}{2K^n} \right) = \frac{1}{(n-1)^K (\log K)^K},
\]
which is summable. By (BC1), eventually for \( K^{n-1} \leq k < K^n \)
\[
S_k \leq \max_{k \leq K^n} S_n \leq c_n = Kh(K^{n-1}) \leq Kh(k).
\]
Hence,
\[ \limsup_k \frac{S_k}{h(k)} \leq K. \]

**Lower bound.** For the other direction, we use a trick similar to the head runs problem. We divide time into big independent blocks. Let \( N > 0 \) large. By the above lemma, for \( \varepsilon > 0 \)
\[
\mathbb{P} \left[ S_{N^n+1} - S_{N^n} > (1 - \varepsilon)h(N^{n+1} - N^n) \right] \geq \frac{1}{\sqrt{2\pi}} \left( y_n^{-1} - y_n^{-3} \right) e^{-y_n^2/2}, \quad (5)
\]
where
\[
y_n = (1 - \varepsilon)\sqrt{2 \log \log(N^{n+1} - N^n)},
\]
so that the RHS in (5) is of order \( n^{-1} \) which sums to \( +\infty \). By (BC2) (by independence) the event in bracket in (5) occurs i.o. Moreover, by the upper bound argument
\[
S_{N^n} > -2h(N^n),
\]
a.s. for large enough \( n \). Hence
\[
S_{N^n+1} > (1 - \varepsilon)h(N^{n+1} - N^n) - 2h(N^n),
\]
for infinitely many \( n \) so that
\[
\limsup_k \frac{S_k}{h(k)} \geq \limsup_n \frac{S_{N^n+1}}{h(N^{n+1})} \geq (1 - \varepsilon)\sqrt{1 - \frac{1}{N}} - 2\sqrt{\frac{1}{N}},
\]
up to logarithmic factors.

**References**


A  Proof of Kronecker’s lemma

Proof: Let \( b_m = \sum_{k=1}^{m} x_k/a_k \). Noting that
\[
x_m = a_m(b_m - b_{m-1}),
\]
we have (setting \( a_0 = b_0 = 0 \))
\[
\frac{1}{a_n} \sum_{m=1}^{n} x_m = \frac{1}{a_n} \left( \sum_{m=1}^{n} a_m b_m - \sum_{m=1}^{n} a_m b_{m-1} \right)
\]
\[
= \frac{1}{a_n} \left( a_n b_n + \sum_{m=1}^{n} a_{m-1} b_{m-1} - \sum_{m=1}^{n} a_m b_{m-1} \right)
\]
\[
= b_n - \sum_{m=1}^{n} \frac{a_m - a_{m-1}}{a_n} b_{m-1}.
\]
Since \( b_n \to b_\infty < +\infty \) and the \( a_n \)'s are non-decreasing, the average on the RHS converges to \( b_\infty \). (Exercise.)

B  St-Petersburg paradox

An important example:

**EX 5.15 (St-Petersburg paradox)** Consider an IID sequence with
\[
\mathbb{P}[X_1 = 2^j] = 2^{-j}, \quad \forall j \geq 1.
\]
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 \( X'_{n,k} = X_{n,k} \mathbb{I}_{|X_{n,k}| \leq b_n} \) 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
\[ \frac{S_n}{n \log_2 n} \to p 1. \]

On the other hand, note that
\[ \mathbb{P}[|X_1| \geq Kn \log_2 n] = \Omega \left( \frac{1}{Kn \log_2 n} \right), \]
which is not summable. By (BC2), since \( K \) is arbitrary,
\[ \limsup_n \frac{S_n}{n \log_2 n} = +\infty, \text{ a.s.} \]

More generally, using the random series results above (see [D]):

**THM 5.16** 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 \( = +\infty \).