Igor Kortchemski
Probability theory,
Autumn 2023
ETH

Chapter 3: Sequences and series of independent random variables

Outline:
1) The use of Borel-Cantelli
2) L^4 version of the strong law of large numbers
3) Kolmogorov's two series theorem
4) Kolmogorov's three series theorem
5) Strong law of large numbers
6) Different notions of convergence
7) Existence of an iid sequence of r.v.

The main goal of this chapter is to study limits of $X_1+\cdots+X_n$ as $n\to\infty$ for independent r.v.
Recall that a property $P=P(\omega)$ is said to be almost sure (a.s.) if $\mathbb{P}(\{\omega\in\Omega : P(\omega) \text{ holds}\}) = \mathbb{P}(P \text{ holds}) = 1$ or, equivalently, $\mathbb{P}(\{\omega\in\Omega: P(\omega) \text{ does not hold}\}) = \mathbb{P}(P \text{ does not hold}) = 0$.

1) The use of Borel-Cantelli

Let $(X_n)_{n\geq 1}$ be a sequence of $\perp\!\!\!\perp$ real valued random variables and $(a_n)_{n\geq 1}$ a given sequence of numbers.
Then by Borel-Cantelli lemmas:

- $\sum_{n=1}^{\infty} \mathbb{P}(X_n > a_n) < \infty \Rightarrow$ a.s there exists n_0 (random) s.t $n\geq n_0 \Rightarrow X_n \leq a_n$
- $\sum_{n=1}^{\infty} \mathbb{P}(X_n > a_n) = \infty \Rightarrow$ a.s $X_n > a_n$ for infinitely many n

(since the events $\{X_n > a_n\}$ are independent)

This is very often used in conjunction with the following lemma.

Lemma Let X_n, X be real r.v. Assume that $\forall \varepsilon > 0$, $\sum_{n=1}^{\infty} \mathbb{P}(|X_n - X| > \varepsilon) < \infty$.
Then a.s $X_n \xrightarrow[n\to\infty]{} X$

Proof: Fix $\varepsilon > 0$. By Borel-Cantelli 1, a.s $|X_n - X| \leq \varepsilon$ for n sufficiently large

⚠ it is not possible to conclude directly that a.s. $\forall \varepsilon > 0$ $|X_n - X| \leq \varepsilon$ for n sufficiently large: in general it is not possible to exchange "a.s" and "$\forall$ on an uncountable set" ⚠

But we can exchange "a.s" and "$\forall$ on a countable set" because a countable intersection of events with probability 1 has probability 1.

The idea is to restrict the values of ε along a (countable) sequence tending to 0:

$\forall k \geq 1$, a.s $|X_n - X| \leq \frac{1}{2^k}$ for n sufficiently large

Thus a.s. $\underbrace{\forall k \geq 1,\ |X_n - X| \leq \frac{1}{2^k} \text{ for } n \text{ sufficiently large}}_{\text{this is equal to the event } \{X_n \xrightarrow[n\to\infty]{} X\}}$

Thus $\mathbb{P}(X_n \to X) = 1$.

END OF LECTURE 10

Corollary Let $(X_n)_{n\geq 1}$ be a sequence of independent identically distributed (iid) real random variables

① If $\mathbb{E}[|X_1|] < \infty$, then a.s $\frac{X_n}{n} \to 0$

② If $\mathbb{E}[|X_1|] = \infty$, then a.s $\frac{X_n}{n}$ does not tend to 0.

③ If $\frac{X_1 + \cdots + X_n}{n}$ converges a.s, then $\mathbb{E}[|X_1|] < \infty$

Proof ① We show that $\forall \varepsilon > 0$, $\sum_{n=1}^{\infty} \mathbb{P}(|\frac{X_n}{n}| > \varepsilon) < \infty$. $(\#)$

To this, using a result from the exercise sheet,

$$\infty > \mathbb{E}\left[\frac{|X_1|}{\varepsilon}\right] = \mathbb{E}\left[\frac{|X_1|}{\varepsilon}\right] = \int_0^{\infty} \mathbb{P}\left(\frac{|X_1|}{\varepsilon} \geq x\right) dx = \sum_{n=0}^{\infty} \int_n^{n+1} \mathbb{P}\left(\frac{|X_1|}{\varepsilon} \geq x\right) dx \geq \sum_{n=0}^{\infty} \mathbb{P}\left(\frac{|X_1|}{\varepsilon} \geq n+1\right)$$

and $(\#)$ follows

② Similarly we show that $\infty = \mathbb{E}[|X_1|] \leq \sum_{n=0}^{\infty} \mathbb{P}(|X_1| \geq n)$,

so since the events $\{|X_n| \geq n\}$ are independent it follows by Borel-Cantelli 2 that a.s $\left|\frac{X_n}{n}\right| \geq 1$ infinitely often.

So a.s $\frac{X_n}{n} \not\to 0$.

③ Set $S_n = X_1 + \cdots + X_n$. Since $n \sim n+1$ as $n \to \infty$, a.s $\frac{S_n}{n} - \frac{S_{n+1}}{n} \longrightarrow 0$.
thus $\frac{X_n}{n} \xrightarrow{a.s} 0$ Thus $\mathbb{E}[|X_1|] < \infty$ by ①.

Our goal is now to go towards the proof of one of the most important results in probability theory

Theorem (Strong law of law large numbers – SLN)
Let $(X_i)_{i \geq 1}$ be an iid sequence of integrable real-valued random variables ($\mathbb{E}[|X_1|] < \infty$)
Then $\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{a.s} \mathbb{E}[X_1]$

Remark . by the previous corollary ③ the integrability condition cannot be removed if we want a finite limit

We start with some variants which are simpler to establish

2) L^4 version of the SLN

Theorem (L^4 version of SLN)
Let $(X_i)_{i \geq 1}$ be an iid sequence of real-valued random variables in L^4 ($\mathbb{E}[X_1^4] < \infty$)
Then $\frac{X_1 + \cdots + X_n}{n} \xrightarrow[n\to\infty]{a.s} \mathbb{E}[X_1]$

Proof of the theorem Without loss of generality it suffices to prove the result when $\mathbb{E}[X_1] = 0$
(for the general case one can then apply the result with $\overline{X}_n = X_n - \mathbb{E}[X_n]$)
Set $S_n = X_1 + \cdots + X_n$ and $K = \mathbb{E}[X_1^4]$
We show that $\sum_{n \geq 1} \mathbb{E}\left[\left(\frac{S_n}{n}\right)^4\right] < \infty$. (*)
Indeed, this will imply by Fubini that $\mathbb{E}\left[\sum_{n \geq 1} \left(\frac{S_n}{n}\right)^4\right] < \infty$, so that a.s $\sum_{n \geq 1} \left(\frac{S_n}{n}\right)^4 < \infty$
Since the general term of a convergent series tends to 0, this indeed implies that a.s. $\frac{S_n}{n} \longrightarrow 0$

To show (*), observe that $S_n^4 = \sum_{1\le j_1,j_2,j_3,j_4\le n} X_{j_1} X_{j_2} X_{j_3} X_{j_4}$, so that
$\mathbb{E}[S_n^4] = \sum_{1\le j_1,j_2,j_3,j_4\le n} \mathbb{E}[X_{j_1} X_{j_2} X_{j_3} X_{j_4}]$

Now, using the $\perp\!\!\!\perp$ and $\mathbb{E}[X_i]=0$, we see that $\mathbb{E}[X_{j_1} X_{j_2} X_{j_3} X_{j_4}]=0$ as soon as one of the indices j_i is different from the other three ones. For example,
$\mathbb{E}[X_1 X_2 X_3 X_2] = \mathbb{E}[X_1]\,\mathbb{E}[X_2 X_3 X_2] = 0\cdot\mathbb{E}[X_2X_3X_2]=0$

Thus, keeping only the remaining terms, we get that

$$\mathbb{E}[S_n^4] = \sum_{j=1}^{n} \mathbb{E}[X_j^4] + 6\sum_{1\le j<j'\le n} \mathbb{E}[(X_j)^2(X_{j'}^2)]$$

Indeed, for each $j<j'$ there are six ways to choose j_1, j_2, j_3, j_4 so that j and j' appear twice each.

Hence $\mathbb{E}[(S_n)^4] = n\,\mathbb{E}[X_1^4] + 3n(n-1)\underbrace{\mathbb{E}[(X_1)^2]^2}_{\le\, \mathbb{E}[X_1^4]=K \text{ by Cauchy Schwarz}}$

We conclude that $\mathbb{E}\left[\left(\frac{S_n}{n}\right)^4\right] \le \frac{3K}{n^2}$ and (*) follows

<u>Application</u> Let $(A_i)_{i\ge1}$ be independent events having same probability p. Then

$$\frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{A_i} \xrightarrow[n\to\infty]{a.s} p$$

This comes from the fact that $(\mathbb{1}_{A_i})_{i\ge1}$ are iid Bernoulli(p) random variables.
This result makes a connection between our "modern" axiomatic approach and the historical definition of probabilities as the frequency of an event happening when repeating an experiment a large number of times.

3) Kolmogorov two series theorem

The Kolmogorov series theorems give conditions for a.s convergence of sums of $\perp\!\!\!\perp$ r.v. We start with a lemma.

<u>Lemma (Kolmogorov maximal inequality)</u> Let $(Z_k)_{1\le k\le n}$ be independent real-valued r.v. in L^2, centred ($\mathbb{E}[Z_k]=0$ for $1\le k\le n$). Set $S_k = Z_1+\dots+Z_k$ for $1\le k\le n$. Then

$$\forall \lambda>0,\quad \mathbb{P}\left(\max_{1\le k\le n}|S_k|\ge\lambda\right) \le \frac{\mathbb{E}[S_n^2]}{\lambda^2}.$$

Remarks • If $X_1, .., X_n$ are ⫫, we have $\mathbb{P}(\max(X_1,..,X_n) \geq \lambda) = 1 - \mathbb{P}(\max(X_1,..,X_n) < \lambda) = 1 - \mathbb{P}(X_1 < \lambda, .., X_n < \lambda)$
which is $1 - \mathbb{P}(X_1 < \lambda) \cdots \mathbb{P}(X_n < \lambda)$ by ⫫.
However, here $S_1, .., S_k$ are in general not ⫫.
• By applying Markov's inequality we get $\mathbb{P}(|S_n| \geq \lambda) = \mathbb{P}(S_n^2 \geq \lambda^2) \leq \frac{1}{\lambda^2} \mathbb{E}[S_n^2]$:
the lemma gives a better inequality since $\max_{1 \leq i \leq n} |S_i| \geq |S_n|$.

Proof Idea: for $1 \leq k \leq n$, set $A_k = \{ |S_k| \geq \lambda, \ |S_i| < \lambda \text{ for } 1 \leq i \leq k-1 \}$.
These events are disjoint and their union in $\{ \max_{1 \leq k \leq n} |S_k| \geq \lambda \}$.
Since they are disjoint, $0 \leq \sum_{k=1}^{n} \mathbb{1}_{A_k} \leq 1$.

Thus $\mathbb{E}[S_n^2] \geq \sum_{k=1}^{n} \mathbb{E}[S_n^2 \mathbb{1}_{A_k}]$. Idea: $S_n^2 = S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2$.

Hence
$$\mathbb{E}[S_n^2] \geq \sum_{k=1}^{n} \mathbb{E}[(S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2) \mathbb{1}_{A_k}]$$
$$\geq \sum_{k=1}^{n} \mathbb{E}[S_k^2 \mathbb{1}_{A_k}] + 2 \sum_{k=1}^{n} \mathbb{E}[S_k (S_n - S_k) \mathbb{1}_{A_k}].$$

But $S_k \mathbb{1}_{A_k}$ is $\sigma(Z_1, .., Z_k)$ measurable and $S_n - S_k$ is $\sigma(Z_{k+1}, .., Z_n)$ measurable
thus they are independent and $\mathbb{E}[S_k \mathbb{1}_{A_k} (S_n - S_k)] = \mathbb{E}[S_k \mathbb{1}_{A_k}] \underbrace{\mathbb{E}[S_n - S_k]}_{=0} = 0$
But
$$\sum_{k=1}^{n} \mathbb{E}[S_k^2 \mathbb{1}_{A_k}] \geq \sum_{k=1}^{n} \mathbb{E}[\lambda^2 \mathbb{1}_{A_k}] = \lambda^2 \sum_{k=1}^{n} \mathbb{P}(A_k)$$
$$= \lambda^2 \mathbb{P}\Big(\bigcup_{k=1}^{n} A_k\Big) = \lambda^2 \mathbb{P}\Big(\max_{1 \leq k \leq n} |S_k| \geq \lambda\Big).$$
Thus $\mathbb{E}[S_n^2] \geq \lambda^2 \mathbb{P}(\max_{1 \leq k \leq n} |S_k| \geq \lambda)$.

END OF LECTURE 11

Theorem (Kolmogorov two series) Let $(Z_n)_{n \geq 1}$ be independent real-valued r.v. in L^2. Assume that
① $\sum_{n \geq 1} \mathbb{E}[Z_n]$ converges
② $\sum_{n \geq 1} \mathrm{Var}(Z_n) < \infty$
Then $\sum_{k=1} Z_k$ converges a.s to a finite r.v.

Remark Here the $(Z_n)_{n\geq 1}$ are not assumed to have same law. (if they have same law and are not constant $=0$, then $\sum_{k\geq 1} \mathrm{Var}(Z_k)=\infty$)

Proof of the theorem Since $\mathrm{Var}(Z_n - \mathbb{E}[Z_n]) = \mathrm{Var}(Z_n)$, we can assume that $\mathbb{E}[Z_n]=0$ (We then apply the result with $Z_n - \mathbb{E}[Z_n]$: we then get $\sum_n (Z_n - \mathbb{E}[Z_n])$

Set $S_n = Z_1 + \dots + Z_n$. The idea is to show that $\forall k \geq 1$, a.s. $\exists m \geq 1$ s.t $\forall n \geq m$, $|S_n - S_m| \leq \frac{1}{k}$ $(*)$

Indeed, by interchanging $\forall \in$ countable set and a.s.

this implies a.s. $\forall k \geq 1$, $\exists m \geq 1$ s.t $\forall n \geq m$, $|S_n - S_m| \leq \frac{1}{k}$. Thus a.s $(S_n)_{n\geq 1}$ is a Cauchy sequence in $\mathbb{R}$, so a.s. it convergences.

To show $(*)$, fix $k \geq 1$ and set $A_m = \{\forall n \geq m : |S_n - S_m| \leq \frac{1}{k}\}$.

We want to show that $\mathbb{P}(\bigcup_{m\geq 1} A_m) = 1$.

But $(A_m)_{m\geq 1}$ is increasing for the inclusion, so $\mathbb{P}(\bigcup_{m\geq 1} A_m) = \lim_{m\to\infty} \mathbb{P}(A_m)$

So we want to show $\mathbb{P}(A_m) \xrightarrow[m\to\infty]{} 1$.

But $1 - \mathbb{P}(A_m) = \mathbb{P}(\exists n \geq m : |S_n - S_m| > \frac{1}{k})$

$$= \lim_{\ell\to\infty} \mathbb{P}(\exists n \text{ with } m \leq n \leq \ell : |S_n - S_m| > \tfrac{1}{k})$$

But by the maximal inequality:

$$\mathbb{P}(\exists n \text{ with } m \leq n \leq \ell : |Z_{m+1} + \dots + Z_n| > \tfrac{1}{k}) \leq k^2\, \mathbb{E}[(Z_{m+1} + \dots + Z_\ell)^2]$$

$$= k^2 (\mathbb{E}[Z_{m+1}^2] + \dots + \mathbb{E}[Z_\ell^2]) \text{ by } \perp\!\!\!\perp \text{ and } \mathbb{E}[Z_i]=0$$

Hence $1 - \mathbb{P}(A_m) \leq k^2 \sum_{i=m+1}^{\infty} \mathbb{E}[Z_i^2]$, which tends to 0 as $m\to\infty$ as the remainder of a convergent series.

4) Kolmogorov's three series theorem

Theorem (Kolmogorov three series theorem) Let $(X_n)_{n\geq 1}$ be $\perp\!\!\!\perp$ real-valued r.v. Assume that there exists $a>0$ such that

① $\sum_{k=1}^{\infty} \mathbb{P}(|X_k| > a) < \infty$

② $\sum_{k=1}^{n} \mathbb{E}[X_k \mathbb{1}_{|X_k|\leq a}] < \infty$ converges as $n\to\infty$

③ $\sum_{k=1}^{\infty} \mathrm{Var}(X_k \mathbb{1}_{|X_k|\leq a}) < \infty$

Then a.s $\sum_{k=1}^{n} X_k$ converges as $n\to\infty$

Remark: it is possible to show that (we will not do it here): if a.s $\sum_{k=1}^{n} X_k$ converges as $n\to\infty$, then ①, ② and ③ hold for every $a>0$

In particular:

- if ①, ② or ③ fails for some $a>0$, then $\mathbb{P}\left(\sum_{k=1}^{n} X_k \text{ does not converge}\right)>0$ and thus a.s $\sum_{k=1}^{n} X_k$ diverges by Kolmogorov's 0-1 law
- ①, ② and ③ hold for one $a>0$ $\iff$ ①, ② and ③ hold for every $a>0$

Proof: to simplify notation, set $Y_n = X_n \mathbb{1}_{|X_n|\leq a}$

By ①, a.s, $|X_k|\leq a$ for k sufficiently large by Borel-Cantelli

thus a.s $Y_k = X_k$ for k sufficiently large.

Therefore a.s. $\sum_{k=1}^{n} X_k$ converges iff $\sum_{k=1}^{n} Y_k$ converges as $n\to\infty$

The random variables $(Y_n)_{n\geq1}$ are $\perp\!\!\!\perp$ and in L^2 (because they are bounded)

By ② the series $\sum_{n\geq1}\mathbb{E}[Y_n]$ converges and by ③ the series $\sum_{n\geq1}\mathrm{Var}(Y_n)$ converges

We conclude that $\sum_{k=1}^{n} Y_k$ converges a.s as $n\to\infty$ by Kolmogorov's two series theorem

5) The strong law of large numbers

Theorem Let $(X_i)_{i\geq1}$ be iid random variables with $\mathbb{E}[|X_1|]<\infty$. Then

$$\frac{X_1+\dots+X_n}{n} \xrightarrow[n\to\infty]{a.s} \mathbb{E}[X_1]$$

The setting is quite different from before: here the r.v. have same distribution and we divide their sums.

We will use the following lemma:

Lemma (Kronecker) Let $(x_n)_{n\geq1}$ be a sequence of real numbers such that $\sum_{k=1}^{n}\frac{x_k}{k}$ converges as $n\to\infty$. Then $\frac{x_1+\dots+x_n}{n}\xrightarrow[n\to\infty]{} 0$.

Proof: Set $w_n = \sum_{k=1}^{n} \frac{x_k}{k}$ and assume $w_n \xrightarrow[n\to\infty]{} w$.
By Cesaro's theorem, $\frac{1}{N}\sum_{n=1}^{N} w_n \xrightarrow[N\to\infty]{} w$.

But $\frac{1}{N}\sum_{n=1}^{N} w_n = \frac{1}{N}\sum_{n=1}^{N}\left(\sum_{j=1}^{n}\frac{x_j}{j}\right) = \frac{1}{N}\sum_{n=1}^{N}\sum_{j=1}^{N} \mathbb{1}_{j\leq n}\frac{x_j}{j} = \frac{1}{N}\sum_{j=1}^{N}\sum_{n=1}^{N}\mathbb{1}_{j\leq n}\frac{x_j}{j}$

So $\frac{1}{N}\sum_{n=1}^{N} w_n = \frac{1}{N}\sum_{j=1}^{N}(N-j+1)\frac{x_j}{j} = \frac{N+1}{N}\sum_{j=1}^{N}\frac{x_j}{j} - \frac{1}{N}\sum_{j=1}^{N} x_j$

Therefore $\frac{1}{N}\sum_{j=1}^{N} x_j = \frac{N+1}{N} w_N - \frac{1}{N}\sum_{n=1}^{N} w_n \xrightarrow[n\to\infty]{} w - w = 0.$

Proof of the theorem Without loss of generality we can assume that $\mathbb{E}[X_1]=0$ (we can add a constant to all the X_i's)

By Kronecker's lemma, if $\sum_{k=1}^{n}\frac{X_k}{k}$ converges a.s. as $n\to\infty$, we would get the result. However this is not always the case, which why we need to work a bit.

First, $\sum_{n\geq 1}\mathbb{P}(|X_n|>n) = \sum_{n\geq 1}\mathbb{P}(|X_1|>n) \leq \mathbb{E}[|X_1|]$ (cf page 2)
$< \infty$

Thus by Borel-Cantelli, a.s. $|X_n|\leq n$ for n sufficiently large
Thus it is enough to show that $\frac{X'_1+\dots+X'_n}{n} \xrightarrow{a.s.} 0$ with $X'_n = X_n \mathbb{1}_{|X_n|\leq n}$

By dominated convergence, $\mathbb{E}[X'_n] = \mathbb{E}[X_1 \mathbb{1}_{|X_1|\leq n}] \xrightarrow[n\to\infty]{} \mathbb{E}[X_1]=0$

Thus, in turn, it is enough to show that $\frac{Y'_1+\dots+Y'_n}{n} \xrightarrow[n\to\infty]{a.s.} 0$ with $Y'_i = X'_i - \mathbb{E}[X'_i]$
To this end we use Kronecker's lemma and show that $\sum_{j=1}^{n}\frac{Y'_j}{j}$ converges a.s. as $n\to\infty$ using Kolmogorov's two series theorem
It is thus enough to show that $\sum_{n\geq 1}\mathbb{E}[(\frac{Y'_n}{n})^2]<\infty$.
For this we have to estimate $\mathbb{E}[(Y'_n)^2] = \mathrm{Var}(Y'_n) = \mathrm{Var}(X'_n) \leq \mathbb{E}[X'^2_n] = \mathbb{E}[|X_1|^2\mathbb{1}_{|X_1|\leq n}]$
To simplify notation, let X be a r.v with the same law as X_1
Thus $\mathbb{E}[|Y'_n|^2] \leq \sum_{j=1}^{n}\mathbb{E}[|X|^2\mathbb{1}_{j-1<|X|\leq j}] \leq \sum_{j=1}^{n} j^2\,\mathbb{P}(j-1<|X|\leq j)$

Hence $\sum_{n=1}^{\infty} \mathbb{E}[(\frac{Y_n'}{n})^2]$

$$\leq \sum_{n=1}^{\infty}\sum_{j=1}^{\infty} \mathbb{1}_{j\leq n} \frac{j^2}{n^2} \mathbb{P}(j-1<|X|\leq j) = \sum_{j=1}^{\infty} j^2 \mathbb{P}(j-1<|X|\leq j)\cdot \underbrace{\sum_{n=j}^{\infty}\frac{1}{n^2}}_{\leq \frac{c}{j}}$$

$$\leq c\sum_{j=1}^{\infty} j\,\mathbb{P}(j-1<|X|\leq j) = c\sum_{j=1}^{\infty}\mathbb{E}[j\,\mathbb{1}_{j-1<|X|\leq j}] \leq c\sum_{j=1}^{\infty}\mathbb{E}[(|X|+1)\,\mathbb{1}_{j-1<|X|\leq j}] = c(\mathbb{E}[|X|]+1)<\infty$$

END OF LECTURE 12

6) Different notions of convergence

Let (X_n), X be r.v in $\mathbb{R}^k$. We equip $\mathbb{R}^k$ with any norm $|\cdot|$ (for example the standard Euclidean norm). We have already seen almost-sure convergence:

$$X_n \xrightarrow[n\to\infty]{a.s} X \text{ if } \mathbb{P}(\{\omega\in\Omega : X_n(\omega)\to X(\omega)\}) = 1$$

Definitions

- We say that X_n converges in probability to X and write $X_n \xrightarrow{\mathbb{P}} X$ if
  $\forall \varepsilon>0,\ \mathbb{P}(|X_n-X|\geq\varepsilon)\xrightarrow[n\to\infty]{} 0.$
- When X_n, X are $\mathbb{R}$-valued, we say that X_n converges in L^p to X and write $X_n \xrightarrow{L^p} X$ if $\mathbb{E}[|X_n-X|^p]\xrightarrow[n\to\infty]{} 0$

Remarks

- by monotonicity, for $\varepsilon'>\varepsilon>0$, $\mathbb{P}(|X_n-X|\geq\varepsilon')\leq\mathbb{P}(|X_n-X|>\varepsilon)\leq\mathbb{P}(|X_n-X|\geq\varepsilon)$,
  so $X_n\xrightarrow{\mathbb{P}}X \Leftrightarrow \forall\varepsilon>0$ small enough $\mathbb{P}(|X_n-X|\geq\varepsilon)\xrightarrow[n\to\infty]{}0$

  $\Leftrightarrow \forall\varepsilon>0$ small enough $\mathbb{P}(|X_n-X|>\varepsilon)\xrightarrow[n\to\infty]{}0$
- a.s convergence involves the joint law $(X,X_1,X_2,\dots)$ while L^p and $\mathbb{P}$ convergence involves only the joint law (X_n,X)

Lemma If $X_n\xrightarrow{\mathbb{P}}X$ and $X_n\xrightarrow{\mathbb{P}}Y$, then a.s $X=Y$

Proof Fix $k\geq1$. Since $\{|X_n-X|<\frac{1}{k}\}\cap\{|X_n-Y|<\frac{1}{k}\}\subset\{|X-Y|<\frac{2}{k}\}$ by the triangular inequality, we get

$$\mathbb{P}(|X-Y|\geq\tfrac{2}{k})\leq\mathbb{P}(|X_n-X|\geq\tfrac{1}{k})+\mathbb{P}(|X_n-Y|\geq\tfrac{1}{k})\xrightarrow[n\to\infty]{}0$$

Thus $\forall k \geq 1$, a.s. $|X-Y| \leq \frac{1}{k}$. By interchanging $\forall \in$ countable set and a.s. we get a.s $\forall k \geq 1$ $|X-Y| \leq \frac{1}{k}$.

Thus a.s $X=Y$.

Proposition $X_n \xrightarrow{\mathbb{P}} X$ iff $\mathbb{E}[\min(|X_n - X|, 1)] \xrightarrow[n\to\infty]{} 0$.

Proof $\boxed{\Rightarrow}$ Take $\varepsilon > 0$ and write

$$\mathbb{E}[\min(|X_n-X|,1)] = \mathbb{E}[\min(|X_n-X|,1)\, \mathbb{1}_{|X_n-X|<\varepsilon}] + \mathbb{E}[\min(|X_n-X|,1)\, \mathbb{1}_{|X_n-X|\geq\varepsilon}]$$
$$\leq \mathbb{E}[\varepsilon\, \mathbb{1}_{|X_n-X|<\varepsilon}] + \mathbb{E}[\mathbb{1}_{|X_n-X|\geq\varepsilon}]$$
$$\leq \varepsilon + \mathbb{P}(|X_n-X|\geq\varepsilon)$$

Thus $\limsup_{n\to\infty} \mathbb{E}[\min(|X_n-X|,1)] \leq \varepsilon$, which implies the desired result

$\boxed{\Leftarrow}$ Write for $\varepsilon \in (0,1]$:

$$\mathbb{P}(|X_n-X|\geq\varepsilon) = \mathbb{P}(\min(|X_n-X|,1)\geq\varepsilon)$$
$$\leq \frac{1}{\varepsilon}\mathbb{E}[\min(|X_n-X|,1)] \quad \text{(Markov inequality)}$$

Proposition If $X_n \xrightarrow{a.s} X$ or $X_n \xrightarrow{L^p} X$ then $X_n \xrightarrow{\mathbb{P}} X$

Proof Assume $X_n \xrightarrow{a.s} X$

Then $\min(|X_n-X|,1) \xrightarrow[n\to\infty]{a.s} 0$ and is dominated by 1, so $\mathbb{E}[\min(|X_n-X|,1)] \xrightarrow{n\to\infty} 0$ by dominated convergence, thus $X_n \xrightarrow{\mathbb{P}} X$ by the previous result.

- Assume $X_n \xrightarrow{L^p} X$. Then $\forall \varepsilon>0$,

$$\mathbb{P}(|X_n-X|\geq\varepsilon) = \mathbb{P}(|X_n-X|^p \geq \varepsilon^p) \leq \frac{1}{\varepsilon^p}\underbrace{\mathbb{E}[|X_n-X|^p]}_{\xrightarrow[n\to\infty]{} 0} \quad \text{(Markov inequality)}$$

Remark In a similar spirit, when $X \in L^2$, writing

$$\mathbb{P}(|X-\mathbb{E}[X]|\geq\varepsilon) \leq \frac{1}{\varepsilon^2}\mathbb{E}[|X-\mathbb{E}[X]|^2] = \frac{1}{\varepsilon^2}\mathbb{E}[(X-\mathbb{E}[X])^2]$$

so

$$\mathbb{P}(|X-\mathbb{E}[X]|\geq\varepsilon) \leq \frac{1}{\varepsilon^2}\mathrm{Var}(X).$$

This is the Bienaymé-Tchebychev inequality

Example Fix $\alpha>0$ and let $(X_n)_{n\geq 1}$ be independent random variables such that $\mathbb{P}(X_n=1)=\frac{1}{n^\alpha}$ and $\mathbb{P}(X_n=0)=1-\frac{1}{n^\alpha}$. Then

- $\mathbb{E}[|X_n|^p]=\frac{1}{n^\alpha}\xrightarrow[n\to\infty]{}0$, so $X_n\xrightarrow{L^p}0$ and thus $X_n\xrightarrow{\mathbb{P}}0$.

- For $\alpha>1$, $\sum_{n=1}^{\infty}\mathbb{P}(X_n=1)<\infty$, so by Borel-Cantelli a.s. $X_n=0$ for n sufficiently large, so a.s. $X_n\longrightarrow 0$.
- For $\alpha\in(0,1]$, $\sum_{n=1}^{\infty}\mathbb{P}(X_n=1)=\infty$, $\sum_{n=1}^{\infty}\mathbb{P}(X_n=0)=\infty$. By independence and by Borel-Cantelli, a.s. $X_n=0$ and $X_n=1$ infinitely often. Thus (X_n) diverges a.s.

Proposition (subsubsequence lemma)
We have $X_n\xrightarrow{\mathbb{P}}X$ iff of every subsequence of (X_n) we can extract a subsubsequence which converges a.s. to X

(a subsequence is an increasing function $\mathbb{N}\to\mathbb{N}$)

In other words, $X_n\xrightarrow{\mathbb{P}}X$ iff $\forall$ subsequence φ, $\exists$ subsequence ψ s.t. $X_{\varphi(\psi(n))}\xrightarrow[n\to\infty]{a.s.}X$

Proof $\boxed{\Rightarrow}$ Let φ be a subsequence. Since $X_{\varphi(n)}\xrightarrow{\mathbb{P}}X$, we know that $\mathbb{E}[\min(|X_{\varphi(n)}-X|,1)]\xrightarrow[n\to\infty]{}0$

We can thus find a subsequence ψ such that $\forall n\geq 1$, $\mathbb{E}[\min(|X_{\varphi\circ\psi(n)}-X|,1)]\leq\frac{1}{2^n}$.

Then $\sum_{n\geq 1}\mathbb{E}[\min(|X_{\varphi\circ\psi(n)}-X|,1)]<\infty$

thus $\mathbb{E}[\sum_{n\geq 1}\min(|X_{\varphi\circ\psi(n)}-X|,1)]<\infty$

Thus a.s. $\sum_{n\geq 1}\min(|X_{\varphi\circ\psi(n)}-X|,1)<\infty$

Thus a.s. $\min(|X_{\varphi\circ\psi(n)}-X|,1)\xrightarrow[n\to\infty]{}0$, thus a.s. $X_{\varphi\circ\psi(n)}\xrightarrow[n\to\infty]{}X$.

$\boxed{\Leftarrow}$ We argue by contradiction: if $X_n\not\xrightarrow{\mathbb{P}}X$, then we can find $\varepsilon>0$ and φ subsequence with $\forall n\geq 1$ $\mathbb{E}[\min(|X_{\varphi(n)}-X|,1)]\geq\varepsilon$ $(\ast)$

Let ψ be subsequence s.t. $X_{\varphi\circ\psi(n)}\xrightarrow{a.s.}X$

Then $\mathbb{E}[\min(|X_{\varphi\circ\psi(n)}-X|,1)]\xrightarrow[n\to\infty]{}0$ because a.s. convergence implies convergence in probability. This contradicts $(\ast)$.

Example (flying saucepans) Consider $[0,1]$ with its Borel σ-field and λ = Lebesgue measure
For $k \geq 0$ and $0 \leq j \leq 2^k - 1$, define

$$X_{2^k+j}(\omega) = \mathbb{1}_{\left[\frac{j}{2^k}, \frac{j+1}{2^k}\right]}(\omega) \quad \text{for } \omega \in [0,1]$$

Then $X_n \xrightarrow{\mathbb{P}} 0$ because for $n \geq 1$ and $\varepsilon \in (0,1]$, $\mathbb{P}(|X_n| \geq \varepsilon) \leq \frac{2}{n}$.
But $\forall \omega \in [0,1]$ there exists infinitely many $n \geq 1$ such that $X_n(\omega) = 1$, so $X_n \overset{a.s}{\not\to} 0$

In the previous example, the portion of space where $X_n \neq 0$ become smaller and smaller, but this portion was moving all around.

Example Take again $[0,1]$ with the Borel σ-field and the Lebesgue measure, and set

$$X_n(\omega) = 2^n \, \mathbb{1}_{[0, \frac{1}{2^n}]}(\omega)$$

Then $X_n \xrightarrow{a.s} 0$ but $\mathbb{E}[|X_n - 0|] = 1 \not\to 0$, so $X_n \overset{L^1}{\not\to} 0$.

In the previous example, the portion of space where $X_n \neq 0$ become smaller and smaller, but on this portion the contribution to the integral is non negligeable because of high values (spikes) on it.

END OF LECTURE 13

The probabilistic notion that prevents such spikes is uniform integrability
If $X \in L^1$, then we have $\mathbb{E}[|X| \mathbb{1}_{|X| \geq K}] \xrightarrow[K \to \infty]{} 0$ by dominated convergence
Uniform integrability extends this to a family of random variables:

Definition A family $(X_i)_{i \in I}$ of integrable real-valued random variables is uniformly integrable, in short UI, if $\sup_{i \in I} \mathbb{E}[|X_i| \mathbb{1}_{|X_i| \geq K}] \xrightarrow[K \to \infty]{} 0$.

Examples ① a finite family of L^1 r.v. is UI

② If $Z \geq 0$ is integrable, $\{X : |X| \leq Z\}$ is UI. Indeed, if $|X| \leq Z$, we have

$\mathbb{E}[|X| \mathbb{1}_{|X| \geq K}] \leq \mathbb{E}[Z \mathbb{1}_{Z \geq K}]$

③ If $(X_i)_{i\in I}$ is bounded in L^p, $p>1$ (i.e. $\sup_{i\in I} \mathbb{E}[|X_i|^p] < \infty$), then $(X_i)_{i\in I}$ is UI. Indeed,

if $C = \sup_{i\in I} \mathbb{E}[|X_i|^p] < \infty$, we have $\mathbb{E}[|X_i| \mathbb{1}_{|X_i| \geq K}] = \mathbb{E}\left[\frac{|X_i|}{|X_i|^p} \cdot |X_i|^p \mathbb{1}_{|X_i| \geq K}\right] \leq \frac{1}{K^{p-1}} \cdot C \xrightarrow[K\to\infty]{} 0$

uniformly in i.

Remark By definition, a sequence $(X_n)_{n\geq 1}$ of integrable r.v is UI if $\sup_{n\geq 1} \mathbb{E}[|X_n| \mathbb{1}_{|X_n| \geq K}] \xrightarrow[K\to\infty]{} 0$

By example ① this is equivalent to $\limsup_{n\to\infty} \mathbb{E}[|X_n| \mathbb{1}_{|X_n| \geq K}] \xrightarrow[K\to\infty]{} 0$

Theorem A family $(X_i)_{i\in I}$ of integrable r.v is UI iff

$(X_i)_{i\in I}$ is bounded in L^1 (i.e. $\sup_{i\in I} \mathbb{E}[|X_i|] < \infty$) and $\forall \varepsilon > 0, \exists \delta > 0, \forall$ event A, $\mathbb{P}(A) \leq \delta \Rightarrow \forall i \in I, \mathbb{E}[|X_i| \mathbb{1}_A] \leq \varepsilon$

($\varepsilon - \delta$ condition)

Before the proof we give a very useful consequence.

Corollary If $(X_i)_{i\in I}$ and $(Y_j)_{j\in J}$ are UI families, then $\{X_i + Y_j : i \in I, j \in J\}$ is UI

Proof of the corollary Since $|X_i + Y_j| \leq |X_i| + |Y_j|$ we get that $(X_i + Y_j)_{i\in I, j\in J}$ is bounded in L^1.

Fix $\varepsilon > 0$. Let $\delta > 0$ be s.t $\mathbb{P}(A) \leq \delta \Rightarrow \mathbb{E}[|X_i| \mathbb{1}_A] \leq \frac{\varepsilon}{2}$ and $\mathbb{E}[|Y_j| \mathbb{1}_A] \leq \frac{\varepsilon}{2}$ $\quad \forall i \in I, j \in J$.

Then $\mathbb{E}[|X_i + Y_j| \mathbb{1}_A] \leq \varepsilon$

Proof of the theorem ($\Rightarrow$) Let $K > 0$ be such that $\sup_{i\in I} \mathbb{E}[|X_i| \mathbb{1}_{|X_i| \geq K}] \leq 1$

Then $\forall i \in I$, $\mathbb{E}[|X_i|] = \mathbb{E}[|X_i| \mathbb{1}_{|X_i| < K}] + \mathbb{E}[|X_i| \mathbb{1}_{|X_i| \geq K}] \leq K + 1$.

Thus $(X_i)_i$ is bounded in L^1.

Now fix $\varepsilon > 0$

Let $K_\varepsilon > 0$ be such that $\sup_{i\in I} \mathbb{E}[|X_i| \mathbb{1}_{|X_i| \geq K_\varepsilon}] \leq \varepsilon$

Then for $\mathbb{P}(A) \leq \frac{\varepsilon}{K_\varepsilon}$ we have for $i \in I$: $\mathbb{E}[|X_i| \mathbb{1}_A] = \mathbb{E}[|X_i| \mathbb{1}_A \mathbb{1}_{|X_i| \leq K_\varepsilon}] + \mathbb{E}[|X_i| \mathbb{1}_A \mathbb{1}_{|X_i| > K_\varepsilon}]$

$\leq K_\varepsilon \mathbb{P}(A) + \varepsilon$

$\leq 2\varepsilon$

$\boxed{\Leftarrow}$ Fix $\varepsilon>0, \delta>0$ such that the ε-δ condition holds.
Let $K>0$ be such that $\frac{1}{K}\sup_{i\in I}\mathbb{E}[|X_i|]\leq\delta$. Then by Markov's inequality $\mathbb{P}(|X_i|\geq K)\leq\frac{\mathbb{E}[|X_i|]}{K}\leq\delta$
so $\mathbb{E}[|X_i|\mathbb{1}_{|X_i|\geq K}]\leq\varepsilon$ (take $A=\mathbb{1}_{|X_i|\geq K}$)

It turns out that UI is precisely what bridges the gap between convergence in probability and L^1 convergence:

Theorem Let $(X_n)_{n\geq 1}$ be integrable real-valued random variable and X a real-valued r.v.
Then $X\in L^1$ and $X_n\xrightarrow{L^1}X$ iff $X_n\xrightarrow{\mathbb{P}}X$ and $(X_n)_{n\geq 1}$ is UI

It can be seen as an extension of the dominated convergence theorem.

Proof $\boxed{\Rightarrow}$ We have already seen that L^1 convergence implies convergence in probability
To show that $(X_n)_{n\geq 1}$ is UI by the corollary it suffices to show that $(X_n-X)_{n\geq 1}$ is UI. To do this, fix $\varepsilon>0$ and choose n_0 s.t $n\geq n_0\Rightarrow\mathbb{E}[|X_n-X|]\leq\varepsilon$.
Let K_0 s.t $K\geq K_0\Rightarrow\max_{1\leq i\leq n_0}\mathbb{E}[|X_i-X|\mathbb{1}_{|X_i-X|\geq K}]\leq\varepsilon$. Then $\sup_{n\geq 1}\mathbb{E}[|X_n-X|\mathbb{1}_{|X_n-X|\geq K}]\leq\varepsilon$ for $K\geq K_0$.

$\boxed{\Leftarrow}$ We first check that $X\in L^1$. By $X_n\xrightarrow{\mathbb{P}}X$, there is a subsequence φ such that $X_{\varphi(n)}\xrightarrow{a.s.}X$
Then by Fatou's lemma, $\mathbb{E}[|X|]\leq\liminf_{n\to\infty}\mathbb{E}[|X_{\varphi(n)}|]<\infty$ since UI $\Rightarrow$ bounded in L^1.
Thus $(X_n-X)_{n\geq 1}$ is UI
Let $\varepsilon,\delta>0$ be such that the ε-δ condition holds. For n sufficiently large, $\mathbb{P}(|X_n-X|\geq\varepsilon)\leq\delta$
So $\mathbb{E}[|X_n-X|]\leq\mathbb{E}[|X_n-X|\mathbb{1}_{|X_n-X|\geq\varepsilon}]+\mathbb{E}[|X_n-X|\mathbb{1}_{|X_n-X|<\varepsilon}]$
$\leq 2\varepsilon$

7) Existence of a sequence of iid random variables

The existence of an iid sequence of r.v of a given law on general spaces is a rather delicate question related to the existence of product measures.

In the case of real-valued random variables, it is possible to do it "by hand" using the existence of the Lebesgue measure.

Consider $(\Omega, \mathcal{A}, \mathbb{P}) = ([0,1], \mathcal{B}([0,1]), \lambda)$ with λ = Lebesgue measure. For $\omega \in \Omega$, $n \geq 1$, set $X_n(\omega) = \lfloor 2^n \omega \rfloor - 2\lfloor 2^{n-1}\omega \rfloor$ when $\lfloor x \rfloor = \sup\{n \in \mathbb{Z} : n \leq x\}$ is the integer part of $x \in \mathbb{R}$

Proposition The r.v $(X_n)_{n\geq 1}$ are iid with $\mathbb{P}(X_1 = 0) = \mathbb{P}(X_1 = 1) = \frac{1}{2}$.

Proof: It is not too difficult to check that $X_n(\omega) \in \{0,1\}$ and that $0 \leq \omega - \sum_{k=1}^{n} \frac{X_k(\omega)}{2^k} \leq \frac{1}{2^n}$ so that $\omega = \sum_{k=1}^{\infty} X_k(\omega) \times \frac{1}{2^k}$, so that the $(X_k(\omega))_{k\geq 1}$ are the coefficients of the dyadic expansion of ω.

For $i_1, \dots, i_p \in \{0,1\}$, we remark that $\{X_1 = i_1, \dots, X_p = i_p\} = [\sum_{j=1}^{p} \frac{i_j}{2^j}, \sum_{j=1}^{p} \frac{i_j}{2^j} + \frac{1}{2^p})$

In particular, $\mathbb{P}(X_1 = i_1, \dots, X_p = i_p) = \frac{1}{2^p}$.

By summing over $i_1, \dots, i_{p-1}$ we get $\mathbb{P}(X_p = i_p) = \frac{1}{2}$. Similarly, $\mathbb{P}(X_j = i_j) = \frac{1}{2}$ for $1 \leq j \leq p$.

Then $\mathbb{P}(X_1 = i_1, \dots, X_p = i_p) = \prod_{k=1}^{p} \mathbb{P}(X_k = i_k)$

Now let $\varphi : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ be one-to-one (for example $\varphi(a,b) = 2^a(2b+1)$).

Define $Y_{i,j} = X_{\varphi(i,j)}$ for $i,j \geq 1$. Then $(Y_{i,j})_{i,j\geq 1}$ are iid with the same law as X_1.

Set

$$W_i = \sum_{j=1}^{\infty} \frac{Y_{ij}}{2^j}.$$

Lemma The r.v. $(W_i)_{i\geq 1}$ are iid uniform on $[0,1]$.

Proof First, W_i is $\sigma(Y_{i,j}: j \geq 1)$ measurable as a limit of $\sigma(Y_{i,j}: j\geq 1)$ measurable functions. Thus the r.v $(W_i)_{i\geq 1}$ are ⫫ by an extension of the coalition principle to infinite families which we have already seen.

Next, for $p\geq 1$, $W_i^{(p)} = \sum_{j=1}^{p} \frac{Y_{i,j}}{2^j}$ has the same law as $X^{(p)} = \sum_{n=1}^{p} \frac{X_n}{2^n}$. Then for $f: \mathbb{R} \to \mathbb{R}_+$ continuous with compact support,

$$\mathbb{E}[f(W_i^{(p)})] = \mathbb{E}[f(X^{(p)})]$$

Taking limits as $p\to\infty$, by continuity of f and by dominated convergence, we get $\mathbb{E}[f(W_i)] = \mathbb{E}[f(X)]$ with X uniform on $[0,1]$.

By Exercise 1(3) of Exercise Sheet 4, this implies $W_i \overset{\text{law}}{=} X$

∎

We can now show:

Proposition Let μ be a probability measure on $\mathbb{R}$. There exists an iid sequence $(Z_i)_{i\geq 1}$ of r.v with law μ

Proof: Set $F_\mu(x) = \mu(]-\infty, x])$ for $x\in\mathbb{R}$ and $F_\mu^{-1}(y) = \inf\{x\in\mathbb{R}: F_\mu(x)\geq y\}$ for $y\in[0,1]$. Then, as in the Lebesgue-Stieltjes construction previously seen, the r.v. $Z_i = F_\mu^{-1}(W_i)$ have law μ, and are ⫫ by the composition principle

∎