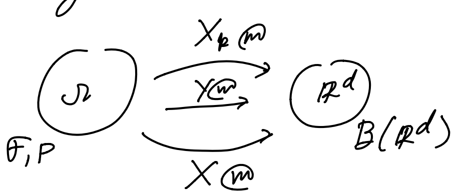
Lecture 17: Convergence a.e. and in probability

We have already seen convergence a.e. and in probability. In this lecture we will collect some useful results used later. Also, we will prove Kolmogorov's SLLN which generalizes the SLLN we proved for bdd r.v.s to quasi-integrable r.v.s.

Assumption for the remainder of this lecture:

Suppose $X, Y, X_1, X_2, \dots$ are random vectors all defined on the same probability space



Definition:

$X_n \xrightarrow{P} X \iff \forall \varepsilon > 0,\ \lim_n P(|X_n - X| \geq \varepsilon) = 0$

$X_n \xrightarrow{a.e.} X \iff P(\lim_n X_n = X) = 1$

Theorem: (uniqueness of limits)

i.e. $P(X=Y)=1$

$X_n \xrightarrow{a.e.} X \ \&\ X_n \xrightarrow{a.e.} Y \implies X \overset{a.e.}{=} Y$

$X_n \xrightarrow{P} X \ \&\ X_n \xrightarrow{P} Y \implies X \overset{a.e.}{=} Y$

Proof:

For $\xrightarrow{a.e.}$ just notice that

$$\underbrace{\{X_n \to X\}} \cap \underbrace{\{Y_n \to Y\}} \subset \{X = Y\}$$

if these have prob 1 then so does this

For $\xrightarrow{P}$ notice that

$$P(|X-Y| \geq \varepsilon) \leq P(|X_n - X| + |X_n - Y| \geq \varepsilon)$$
$$\leq P(|X_n - X| \geq \tfrac{\varepsilon}{2}) + P(|X_n - Y| \geq \tfrac{\varepsilon}{2})$$
$$\longrightarrow 0 \text{ as } \varepsilon \to 0$$

$\therefore\ \{|X-Y| \geq \frac{1}{n}\} \uparrow \{|X-Y| > 0\}$ and by continuity from above we have

$$P(|X-Y|>0) = \lim_n P(|X-Y| \geq \tfrac{1}{n}) = 0$$

QED

Theorem (i.o characterization)

$$X_n \xrightarrow{a.e.} X \iff P(|X_n - X| \geq \varepsilon \text{ i.o. } n) = 0 \text{ for all } \varepsilon > 0$$
$$\iff P(|X_n - X| < \varepsilon \text{ a.a. } n) = 1 \text{ for all } \varepsilon > 0$$

Proof:

We have seen this before but just for completeness:

$$\{X_n \not\to X\} = \bigcup_{\varepsilon \in \mathbb{Q}^+} \{|X_n - X| \geq \varepsilon \text{ i.o. } n\}$$

Now $P(X_n \to X) = 1$

$$\iff P(X_n \not\to X) = 0$$
$$\iff P\Big(\bigcup_{\varepsilon \in \mathbb{Q}^+} \{|X_n - X| \geq \varepsilon \text{ i.o. } n\}\Big) = 0$$
$$\iff P(|X_n - X| \geq \varepsilon \text{ i.o. } n) = 0 \quad \forall \varepsilon \in \mathbb{Q}^+$$

To finish let $\varepsilon' \in \mathbb{R}^+$ & let $\varepsilon < \varepsilon'$ s.t. $\varepsilon \in \mathbb{Q}^+$. Then

$$P(|X_n - X| \geq \varepsilon' \text{ i.o.n}) \leq P(|X_n - X| \geq \varepsilon \text{ i.o.n})$$

so that

$$P(|X_n - X| \geq \varepsilon \text{ i.o.n}) = 0 \quad \forall \varepsilon \in \mathbb{Q}^+$$
$$\Updownarrow$$
$$P(|X_n - X| \geq \varepsilon \text{ i.o.n}) = 0 \quad \forall \varepsilon \in \mathbb{R}^+$$

For the a.a.n criterion take complements.

QED

The above i.o. characterization & Fatou establishes the following result:

Theorem

$$X_n \xrightarrow{a.e.} X \Rightarrow X_n \xrightarrow{P} X.$$

Here is one condition which gives the reverse implication

Theorem:

If the X_n's are real valued and if for P-a.e. $\omega \in \Omega$ $X_n(\omega)$ is monotonic in n then

$$X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{a.e.} X.$$

Proof:

By monotonicity $\exists$ a r.v. Y s.t.

$$X_n \xrightarrow{a.e.} Y \text{ — possibly extended}$$

$\therefore$ $X_n \xrightarrow{P} Y = X$ by uniqueness of limits

$\therefore$ $X_n \xrightarrow{a.e.} X$. QED

Theorem (Cauchy criterion)

$$\exists X \text{ s.t. } X_n \xrightarrow{a.e.} X$$
$$\Updownarrow$$
$$\forall \varepsilon > 0, \ \lim_n \lim_m P\Big(\max_{n \leq p \leq m} |X_n - X_p| \geq \varepsilon\Big) = 0$$

and

$$\exists X \text{ s.t. } X_n \xrightarrow{P} X$$
$$\Updownarrow (*)$$
$$\forall \varepsilon > 0, \ \lim_n \lim_m \max_{n \leq p \leq m} P(|X_n - X_p| \geq \varepsilon) = 0$$
$$\Updownarrow (*)$$
$$\forall \varepsilon > 0, \ \lim_n \lim_m P(|X_n - X_m| \geq \varepsilon) = 0$$

Proof:

Recall that we already proved the condition for $\xrightarrow{a.e.}$ in Lecture 15 as we needed it for Kolmogorov's 3 series thm.

For $\overset{(*)}{\Longrightarrow}$: Suppose $X_n \xrightarrow{P} X$

$$\therefore \lim_{n\, m} \max_{n \leq p \leq m} P(|X_n - X_p| \geq \varepsilon)$$
$$\leq \lim_{n\, m} \max_{n \leq p \leq m} \Big[P(|X_n - X| \geq \tfrac{\varepsilon}{2}) + P(|X - X_p| \geq \tfrac{\varepsilon}{2})\Big]$$
$$\leq \underbrace{\lim_n P(|X_n - X| \geq \tfrac{\varepsilon}{2})}_{=0} + \underbrace{\underbrace{\lim_n \sup_{n \leq p}}_{\equiv \limsup_n} P(|X - X_p| \geq \tfrac{\varepsilon}{2})}_{=0}$$

For $\overset{(*)}{\Longleftarrow}$: Let ε_k, ρ_k be positive and summable in k. Now inductively choose $n_k > n_{k-1}$ s.t.

$$\lim_m \max_{n_k \leq p \leq m} P(|X_{n_k} - X_p| \geq \varepsilon_k) \leq \rho_k$$

$$\therefore P(|X_{n_k} - X_{n_{k+1}}| \geq \varepsilon_k) \leq \rho_k \quad (*)$$

5

Since $\sum_k p_k < \infty$ the FBC Lemma implies

$$P\left(|X_{n_k} - X_{n_{k+1}}| \geq \varepsilon_k \text{ i.o. } k\right) = 0$$

$$\therefore P\left(|X_{n_k} - X_{n_{k+1}}| < \varepsilon_k \text{ a.a. } k\right) = 1$$

Notice that if a sequence of numbers x_k satisfies $|x_k - x_{k+1}| < \varepsilon_k$ $\forall$ suff large k then

$$\lim_{N\to\infty} \sup_{N\leq n\leq m} |x_n - x_m| \leq \lim_{N\to\infty} \sup_{N\leq n\leq m} \sum_{k=n}^{m-1} |x_k - x_{k+1}|$$

$$\leq \underbrace{\lim_{N\to\infty} \sup_{N\leq n\leq m} \sum_{k=n}^{m-1} \varepsilon_k}_{\text{This is zero since } \sum_{k=1}^{\infty} \varepsilon_k < \infty}$$

i.e. $\{x_k\}_{k\geq 1}$ is a Cauchy sequence.

$\therefore$ For P-a.e. $\omega\in\Omega$ $\{X_k(\omega)\}_{k\geq 1}$ is Cauchy

$\therefore \exists$ r.v. X s.t. $X_{n_k} \xrightarrow{a.e.} X$ as $k\to\infty$

must take values in $\mathbb{R}$ P-a.e. since $\mathbb{R}$ is complete.

$\therefore X_{n_k} \xrightarrow{P} X$ as $k\to\infty$

Now $X_n \xrightarrow{P} X$ since

$$P(|X_n - X| \geq \varepsilon) \leq P(|X_n - X_{n_k}| \geq \tfrac{\varepsilon}{2}) + \underbrace{P(|X_{n_k} - X| \geq \tfrac{\varepsilon}{2})}_{\to 0 \text{ as } k\to\infty}$$

implies that

6

$$\limsup_k \limsup_n P(|X_n - X| \geq \varepsilon)$$

$$\leq \limsup_k \limsup_n P(|X_n - X_{n_k}| \geq \tfrac{\varepsilon}{2})$$

$$\leq \underbrace{\limsup_m \limsup_n \underbrace{P(|X_n - X_m| \geq \tfrac{\varepsilon}{2})}_{\leq \max_{m\leq p\leq n} P(|X_m - X_p| \geq \frac{\varepsilon}{2})}}_{=0 \text{ by assumption.}}$$

For $\overset{(*)}{\Longleftarrow}$ The argument is similar. The key for $\Leftarrow$ is to just show $(*)$ holds.

QED.

The next theorem is analogous to a mix of the sub-sub-seq test for convergence and the skorkhod representation Thm. It's usage is also similar in that it allows one to extend integration thms that require $\xrightarrow{a.e.}$ to the weaker condition $\xrightarrow{P}$.

... and gives a continuous mapping theorem for $\xrightarrow{P}$.

Theorem: (Sub-sub-seq for $\xrightarrow{P}$)

$X_n \xrightarrow{P} X$ iff $\forall$ sub-seq n_k $\exists$ sub-sub-seq n_{k_j} s.t. $X_{n_{k_j}} \xrightarrow{a.e.} X$ as $j\to\infty$.

Proof:

$(\Longrightarrow)$ For a given n_k recursively choose $n_{k_j} > n_{k_{j-1}}$ so that

$$\sum_{j=1}^{\infty} P\left(|X_{n_{k_j}} - X| \geq \tfrac{1}{j}\right) < \infty$$

7

By the FBC Lemma

$$P(|X_{n_{k_j}} - X| \geq \tfrac{1}{j} \text{ i.o. } j) = 0$$

$\therefore$ $P(|X_{n_{k_j}} - X| \geq \tfrac{1}{j} \text{ a.a. } j) = 1$

$\therefore$ $X_{n_{k_j}} \xrightarrow{a.e.} X$ as $j \to \infty$.

($\Longleftarrow$) Argue by contradiction.

$X_n \overset{P}{\not\to} X \Rightarrow \exists \varepsilon > 0$ s.t. $P(|X_n - X| \geq \varepsilon) \not\to 0$

$\Rightarrow \exists \varepsilon, \delta > 0$ s.t. $P(|X_{n_k} - X| \geq \varepsilon) \geq \delta$ $\forall k$ where n_k is some sub-seq.

But by assumption $\exists n_{k_j}$ s.t.

$$X_{n_{k_j}} \xrightarrow{a.e.} X$$

$\therefore$ $X_{n_{k_j}} \xrightarrow{P} X$

$\therefore$ $P(|X_{n_{k_j}} - X| \geq \varepsilon) \to 0$ as $j \to \infty$

contradictory

QED

Lets first use the above result for a continuous mapping theorem.

Theorem: (continuous mapping)

If $g: \mathbb{R}^d \to \mathbb{R}^k$ is X-continuous (i.e. $PX^{-1}(\underbrace{\{x \in \mathbb{R}^d : g \text{ is continuous @ } x\}}_{:= C_g}) = 1$)

Then

$$X_n \xrightarrow{a.e.} X \Rightarrow g(X_n) \xrightarrow{a.e.} g(X)$$

and $X_n \xrightarrow{P} X \Rightarrow g(X_n) \xrightarrow{P} g(X)$.

Proof:

For $\xrightarrow{a.e.}$ just notice that

$$\underbrace{\{X_n \to X\} \cap \{X \in C_g\}}_{\text{has probability } 1} \subset \{g(X_n) \to g(X)\}$$

8

For $\xrightarrow{P}$ notice that

$X_n \xrightarrow{P} X \Leftrightarrow \forall n_k \exists n_{k_j}$ s.t. $X_{n_{k_j}} \xrightarrow{a.e.} X$ by the sub-sub-seq thm

$\Longrightarrow \forall n_k \exists n_{k_j}$ s.t. $g(X_{n_{k_j}}) \xrightarrow{a.e.} g(X)$ by continuous mapping for $\xrightarrow{a.e}$

$\Leftrightarrow g(X_n) \xrightarrow{P} g(X)$ again by the sub-sub-seq thm

QED

Now lets see how to use the sub-sub-seq thm to show the sandwich thm for $\xrightarrow{P}$.

Theorem: ($\xrightarrow{P}$ sandwich)

If (i) $0 \leq X_n \leq Y_n$ P-a.e.

$$\begin{array}{cc} X_n & Y_n \\ \downarrow P & \downarrow P \\ X & Y \end{array}$$

(ii) $Y_n, Y \in L_1(\Omega, \mathcal{F}, P)$ & $E(Y_n) \to E(Y)$

then $X_n, X \in L_1(\Omega, \mathcal{F}, P)$ & $E(X_n) \to E(X)$.

Proof:

Lets first show the $\xrightarrow{a.e.}$ version by assuming $X_n \xrightarrow{a.e.} X$ & $Y_n \xrightarrow{a.e.} Y$.

Then

$$\begin{aligned} E(X) = E(\liminf_n X_n) &\leq \liminf_n E(X_n), \quad \text{Fatou} \\ &\leq \liminf_n E(Y_n), \quad \text{Big 3} \\ &= E(Y) < \infty \end{aligned}$$

$\therefore$ $X, X_n \in L_1(\Omega, \mathcal{F}, P)$.

Also $0 \leq Y_n - X_n$ implies

$$\begin{aligned} E(Y) - E(X) = E(Y - X) &\quad \text{Big 3} \\ &= E \liminf_n (Y_n - X_n) \\ &\leq \liminf (EY_n - EX_n), \quad \text{Fatou} \\ &= E(Y) - \limsup_n EX_n, \quad \text{assump} \end{aligned}$$

Since $E(Y) < \infty$ we can subtract from both sides to obtain

$$E(X) \le \liminf_n E(X_n) = \limsup_n E(X_n) \le E(X)$$

so, indeed, $E(X_n) \to E(X)$.

Now replace the assumption $X_n \xrightarrow{a.e.} X$ & $Y_n \xrightarrow{a.e.} Y$ with $X_n \xrightarrow{P} X$ & $Y_n \xrightarrow{P} Y$.

Proceed by contradiction and suppose

$$E(X_n) \not\to E(X).$$

$\therefore \exists n_k$ & $\delta > 0$ s.t.

$$(*) \qquad |E(X_{n_k}) - E(X)| \ge \delta \quad \forall k$$

But the sub-sub-seq Thm implies $\exists n_{k_j}$ s.t.

$$X_{n_{k_j}} \xrightarrow{a.e.} X \qquad \&$$

$$Y_{n_{k_j}} \xrightarrow{a.e.} Y$$

Now $E(X_{n_{k_j}}) \to E(X)$ by a.e. result. This contradicts $(*)$ & so

$$E(X_n) \to E(X)$$

QED

Stochastic Order notation

If X_n's and Y_n's are r.v.s we write

$X_n = o_p(Y_n)$ to mean $\frac{X_n}{Y_n} \xrightarrow{P} 0$

$X_n = O_p(Y_n)$ to mean $\left\{ \frac{X_n}{Y_n} \right\}_{n \ge 1}$ is tight

In particular

$E|X_n|^p = O(1)$ some $p > 0$ $\Rightarrow$ $X_n = O_p(1)$

$\overset{\text{Prohorov}}{\Rightarrow} \exists n_k$ and a r.v. X s.t.

s.t. $X_{n_k} \xrightarrow{\mathcal{D}} X$

Kolmogorov's SLLN

In Lecture 11 we proved the following SLLN for bdd r.v.s:

Theorem: (SLLN for bounded r.v.s)

Let $X_1, X_2, \ldots$ be iid r.v.s on $(\Omega, \mathcal{F}, P)$ s.t. $|X_i| \le c$ for some finite c. Then

$$\frac{S_n}{n} \xrightarrow{a.e.} E(X_1)$$

where $S_n = X_1 + \cdots + X_n$.

The most general result is called Kolmogorov's SLLN an says

Theorem: (Kolmogorov's SLLN)

Let $X_1, X_2, \ldots$ be iid r.v.s on $(\Omega, \mathcal{F}, P)$ which are quasi-integrable. Then

$$\frac{S_n}{n} \xrightarrow{a.e.} E(X_1)$$

where $S_n = X_1 + \cdots + X_n$.

The proof of this theorem is attached below for your refenece. As warmup we show a weaker version for L_2 r.v.s.

Theorem: (SLLN for L_2 r.v.s)

Let $X_1, X_2, \ldots$ be iid r.v.s on $(\Omega, \mathcal{F}, P)$ s.t. $E(X_1^2) < \infty$. Then

$$\frac{S_n}{n} \xrightarrow{a.e.} E(X_1)$$

where $S_n = X_1 + \cdots + X_n$.

Proof:

Notice that w.l.g we can assume $X \geq 0$ since

if $\quad \frac{S_{n,+}}{n} := \frac{X_1^+ + \cdots + X_n^+}{n} \xrightarrow{a.e.} E(X_1^+)$

and $\quad \frac{S_{n,-}}{n} := \frac{X_1^- + \cdots + X_n^-}{n} \xrightarrow{a.e.} E(X_1^-)$

Then $\quad \frac{S_n}{n} = \frac{S_{n,+}}{n} - \frac{S_{n,-}}{n} \xrightarrow{a.e.} \underbrace{E(X_1^+) - E(X_1^-)}_{= E(X_1)}$

$\therefore$ we can assume $X \geq 0$.

By Chebyshev

$$P\left(\left|\frac{S_n}{n} - E\left(\frac{S_n}{n}\right)\right| \geq \varepsilon\right) \leq \frac{\operatorname{var}(S_n/n)}{\varepsilon^2} \leq \frac{E(X_1^2)}{\varepsilon^2 n}$$

Consider a sub-seq $n_k = \lceil \alpha^k \rceil$ where $\alpha \in (1, \infty)$.
Since $\sum_{k=1}^{\infty} \frac{E(X_1^2)}{n_k^2 \varepsilon} < \infty$ the FBC lemma gives

$$P\left(\left|\frac{S_{n_k}}{n_k} - \underbrace{E\left(\frac{S_{n_k}}{n_k}\right)}_{= E(X_1)}\right| \geq \varepsilon \text{ i.o. } k\right) = 0$$

So $\quad \frac{S_{n_k}}{n_k} \xrightarrow{a.e.} E(X_1)$ as $k \to \infty$.

Now we use the fact that $X \geq 0$ to show $\frac{S_n}{n} \xrightarrow{a.e.} E(X_1)$.

Notice that when $n_k \leq n \leq n_{k+1}$

$$\frac{S_{n_k}}{n_{k+1}} \leq \frac{S_n}{n} \leq \frac{S_{n_{k+1}}}{n_k}$$

Note: n_{k+1} here instead of n_k

13

Since for every $\alpha \in (1, \infty)$

$$LHS = \frac{S_{n_k}}{n_{k+1}} = \frac{S_{n_k}}{n_k}\frac{n_k}{n_{k+1}} \xrightarrow{a.e.} \frac{E(X_1)}{\alpha}$$

$$RHS = \frac{S_{n_{k+1}}}{n_k} = \frac{S_{n_{k+1}}}{n_{k+1}}\frac{n_{k+1}}{n_k} \xrightarrow{a.e.} \alpha E(X_1)$$

as $k \to \infty$ we have

$$P\Bigg[\underbrace{\bigcap_{\substack{\alpha \geq 1 \\ \alpha \in \mathbb{Q}}} \left\{ \frac{E(X_1)}{\alpha} \leq \liminf_n \frac{S_n}{n} \leq \limsup_n \frac{S_n}{n} \leq \alpha E(X_1) \right\}}_{= \left\{ \frac{S_n}{n} \to E(X_1) \right\}}\Bigg] = 1$$

as was to be shown.

QED

14

Theorem 117 (Kolmogorov's SLLN). *Let $X_1, X_2, \ldots$ be independent random variables, each distributed like some random variable X, all defined on the same probabilty space. Let $S_n := X_1 + \ldots + X_n$.*

- *If X is quasi-integrable then $S_n/n \xrightarrow{ae} E(X)$.*

Proof. The main idea is to mimic arguments for Theorem 116 but with an additional truncation argument. Again we can suppose without loss of generality that X is positive.

First consider the case $E(X) < \infty$. The idea is to analyze the truncated average T_n/n instead of S_n/n where

$$T_n := \sum_{i=1}^n X_i I_{\{X_i \le i\}}.$$

Notice that for large i the terms $X_i I_{\{X_i \le i\}}$ start to behave more like X_i. Moreover the small i terms in T_n/n are downweighted by $1/n$. Therefore once might expect T_n/n to behave like S_n/n for large n. To continue the proof we again we use Chebyshev

$$\begin{aligned} P\Big[\big|T_n/n - E(T_n/n)\big| \ge \epsilon\Big] &\le \frac{\mathrm{var}(T_n/n)}{\epsilon^2} \\ &\le \frac{1}{\epsilon^2 n^2}\sum_{i=1}^n E(X_i^2 I_{\{X_i\le i\}}) \\ &\le \frac{1}{\epsilon^2 n^2}\sum_{i=1}^n E(X_i^2 I_{\{X_i\le n\}}) \\ &\le \frac{E(X^2 I_{\{X\le n\}})}{\epsilon^2 n}. \end{aligned} \tag{49}$$

We now notice that if we define the subsequence $n_k := \lceil \alpha^k \rceil$ where $\alpha \in (1,\infty)$ then the right hand side (above) is summable. In particular

$$\begin{aligned} \sum_{k=1}^\infty \frac{E(X^2 I_{\{X\le n_k\}})}{n_k} &\underset{\text{Fubini}}{=} E\Big(X^2 \sum_{k=1}^\infty \frac{1}{n_k} I_{\{X\le n_k\}}\Big) \\ &= E\Big(X^2\Big[0 + \cdots 0 + \frac{1}{n_j} + \frac{1}{n_{j+1}} + \cdots\Big]\Big) \end{aligned}$$

where j is the first index such that $X \le n_j$, i.e. $\frac{X}{n_j} \le 1$. Also notice the higher order terms can be bounded as follows

$$\frac{X}{n_{j+m}} = \frac{X}{\lceil \alpha^{j+m}\rceil} \le \frac{X}{\alpha^{j+m}} = \frac{1}{\alpha^m}\frac{n_j}{\alpha^j}\frac{X}{n_j} \le \frac{2}{\alpha^m}.$$

Therefore

$$X^2\Big[\frac{1}{n_j} + \frac{1}{n_{j+1}} + \cdots\Big] \le X\Big[\frac{2}{\alpha^0} + \frac{2}{\alpha^1} + \cdots\Big] \tag{50}$$

Now since $E(X) < \infty$, the right hand side of (50) has finite expected value, and hence Borel-Cantelli gives

$$T_{n_k}/n_k - E(T_{n_k}/n_k) \xrightarrow{ae} 0 \tag{51}$$

as $k \to \infty$. Now if we can show that $E(T_{n_k}/n_k) = \mu + o(1)$ we can apply the same arguments as found in Theorem 116 to get

$$T_n/n \xrightarrow{ae} \mu \tag{52}$$

as $n \to \infty$.

Now we show $E(T_n/n) = \mu + o(1)$ and $T_n/n = S_n/n + o(1)$ with probability one. Notice that $E(T_n/n) = \frac{1}{n}\sum_{i=1}^n E(X_i I_{\{X_i\le i\}})$ where $\lim_i E(X_i I_{\{X_i \le i\}}) = \lim_i E(X I_{\{X\le i\}}) = E(X) = \mu$ by the DCT. Therefore Lemma 7 applies with $\mu_i := E(X_i I_{\{X_i \le i\}})$ to give

$$E(T_n/n) = \frac{1}{n}\sum_{i=1}^n \mu_i = \mu + o(1). \tag{53}$$

To finish lets analyze the terms in T_n versus the terms in S_n

$$P(X_i \ne X_i I_{\{X_i\le i\}}) = P(X_i > i).$$

Lemma 6 (below) gives that $\sum_{i=1}^\infty P(X_i > i) = E(\lceil X \rceil) < \infty$. Borel-Cantelli then gives $P(X_i \ne X_i I_{\{X_i \le i\}}\ i.o._i) = 0$ which implies that for the high-index terms in T_n are eventually exactly the same as in S_n. Therefore

$$T_n/n = S_n/n + o(1) \tag{54}$$

with probability one. Equations (51), (54) and (53) finish the proof of the case when $E(X) < \infty$.

Now consider the case $E(X) = \infty$. We simply show that $\liminf_n S_n/n = \infty$ with probability one (which allows us to conclude that $\liminf_n S_n/n = \limsup_n S_n/n = \lim_n S_n/n = \infty$ with probability one). Indeed

$$\begin{aligned} \liminf_{n\to\infty} \frac{S_n(w)}{n} &\ge \liminf_{n\to\infty} \frac{X_1(w)\wedge k + \cdots + X_n(w)\wedge k}{n} \\ &= E(X\wedge k), \quad \text{by the case above} \end{aligned}$$

for all $w \in A_k$ where $P(A_k) = 1$. Continuity from below in Big 3 implies $E(X \wedge k) \to \infty$. Therefore $\liminf_{n\to\infty} S_n(w)/n = \infty$ for all $w \in \cap_{k=1}^\infty A_k$ which has probability one. Therefore

$$S_n/n \xrightarrow{ae} \infty.$$

□

The following lemma was used in the above proof to analyze the difference between a truncated sum and the non-truncated sum.

Lemma 6 (Expect the ceiling lemma). *If X is a nonnegative random variable, then*

$$\sum_{i=0}^\infty P(X > i) = E(\lceil X \rceil). \tag{55}$$

Proof.

$$\sum_{i=0}^\infty P(X > i) = \sum_{i=0}^\infty E(I_{\{X>i\}}) \underset{\text{Fubini}}{=} E\Big(\underbrace{\sum_{i=0}^\infty I_{\{X>i\}}}_{=\lceil X\rceil}\Big).$$

□

The following lemma was used to show that the expected value of a truncated sum, in the most general proof of the SLLN, converges to the non-truncated expected value.

Lemma 7 (Cesàr summation lemma). *If* $\mu_i \to \mu$ *as* $i \to \infty$, *then* $\left(\sum_{i=1}^{n} \mu_i\right)/n \to \mu$ *as* $n \to \infty$.

Proof.

$$\begin{aligned}
\left|\frac{1}{n}\sum_{i=1}^{n} \mu_i - \mu\right| &\leq \frac{1}{n}\sum_{i=1}^{n} |\mu_i - \mu| \\
&\leq \frac{1}{n}\sum_{i=1}^{m} |\mu_i - \mu| + \sup_{i>m} |\mu_i - \mu|, \ m \leq n \\
&=: I_{n,m} + II_m
\end{aligned}$$

Taking a limit as $n \to \infty$ first one gets $\lim_n I_{n,m} = 0$, then take a limit as $m \to \infty$ to get $\lim_m II_m = \limsup_m |\mu_m - \mu| = 0$. $\square$