Lecture-14: Almost sure convergence

1 Point-wise convergence

Consider a random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ defined on a probability space $(\Omega, \mathcal{F}, P)$, then each $X_n \triangleq \pi_n \circ X : \Omega \to \mathbb{R}$ is a random variable. There are many possible definitions for convergence of a sequence of random variables. One idea is to consider $X(\omega) \in \mathbb{R}^{\mathbb{N}}$ as a real valued sequence for each outcome ω, and consider the $\lim_n X_n(\omega)$ for each outcome ω.

Definition 1.1. A random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ **converges point-wise** to a random variable $X_\infty : \Omega \to \mathbb{R}$, if for all outcomes $\omega \in \Omega$, we have

$$\lim_n X_n(\omega) = X_\infty(\omega).$$

Remark 1. This is a very strong convergence. Intuitively, what happens on an event of probability zero is not important. We will strive for a weaker notion of convergence, where the sequence of random variable converge point-wise on a set of outcomes with probability one.

2 Almost sure statements

Definition 2.1. A statement holds **almost surely** (a.s.) if there exists an event called the **exception set** $N \in \mathcal{F}$ with $P(N) = 0$ such that the statement holds for all $\omega \notin N$.

Example 2.2 (Almost sure equality). Two random variables X, Y defined on the probability space $(\Omega, \mathcal{F}, P)$ are said to be equal a.s. if the following exception set

$$N \triangleq \{\omega \in \Omega : X(\omega) \neq Y(\omega)\} \in \mathcal{F},$$

has probability measure $P(N) = 0$. Then Y is called a **version** of X, and we can define an equivalence class of a.s. equal random variables.

Example 2.3 (Almost sure monotonicity). Two random variables X, Y defined on the probability space $(\Omega, \mathcal{F}, P)$ are said to be $X \leqslant Y$ a.s. if the exception set $N \triangleq \{\omega \in \Omega : X(\omega) > Y(\omega)\} \in \mathcal{F}$ has probability measure $P(N) = 0$.

3 Almost sure convergence

Definition 3.1 (Almost sure convergence). A random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ defined on the probability space $(\Omega, \mathcal{F}, P)$ **converges almost surely**, if the following exception set

$$N \triangleq \left\{\omega \in \Omega : \liminf_n X_n(\omega) < \limsup_n X_n(\omega) \text{ or } \limsup_n X_n(\omega) = \infty\right\} \in \mathcal{F},$$

has zero probability. Let X_∞ be the point-wise limit of the sequence of random variables $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ on the set N^c, then we say that the sequence X **converges almost surely** to X_∞, and denote it as

$$\lim_n X_n = X_\infty \text{ a.s.}$$

Example 3.2 (Convergence almost surely but not everywhere). Consider the probability space $([0,1],\mathcal{B}([0,1]),\lambda)$ such that $\lambda([a,b]) = b-a$ for all $0 \leqslant a \leqslant b \leqslant 1$. For each $n \in \mathbb{N}$, we define the scaled indicator random variable $X_n : \Omega \to \{0,1\}$ such that

$$X_n(\omega) \triangleq n\mathbb{1}_{[0,\frac{1}{n}]}(\omega).$$

Let $N = \{0\}$, then for any $\omega \notin N$, there exists $m = \lceil \frac{1}{\omega} \rceil \in \mathbb{N}$, such that for all $n > m$, we have $X_n(\omega) = 0$. That is, $\lim_n X_n = 0$ a.s. since $\lambda(N) = 0$. However, $X_n(0) = n$ for all $n \in \mathbb{N}$.

4 Convergence in probability

Definition 4.1 (convergence in probability). A random sequence $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ defined on the probability space $(\Omega,\mathcal{F},P)$ **converges in probability** to a random variable $X_\infty : \Omega \to \mathbb{R}$, if $\lim_n P(A_n(\epsilon)) = 0$ for any $\epsilon > 0$, where

$$A_n(\epsilon) \triangleq \{\omega \in \Omega : |X_n(\omega) - X_\infty(\omega)| > \epsilon\} \in \mathcal{F}.$$

Remark 2. $\lim_n X_n = X_\infty$ a.s. means that for almost all outcomes ω, the difference $X_n(\omega) - X_\infty(\omega)$ gets small and stays small.

Remark 3. $\lim_n X_n = X_\infty$ i.p. is a weaker convergence than a.s. convergence, and merely requires that the probability of the difference $X_n(\omega) - X_\infty(\omega)$ being non-trivial becomes small.

Example 4.2 (Convergence in probability but not almost surely). Consider the probability space $([0,1],\mathcal{B}([0,1]),\lambda)$ such that $\lambda([a,b]) = b-a$ for all $0 \leqslant a \leqslant b \leqslant 1$. For each $k \in \mathbb{N}$, we consider the sequence $S_k = \sum_{i=1}^{k} i$, and define integer intervals $I_k \triangleq \{S_{k-1}+1,\dots,S_k\}$. Clearly, the intervals $(I_k : k \in \mathbb{N})$ partition the natural numbers, and each $n \in \mathbb{N}$ lies in some I_{k_n}, such that $n = S_{k_n-1} + i_n$ for $i_n \in [k_n]$. Therefore, for each $n \in \mathbb{N}$, we define indicator random variable $X_n : \Omega \to \{0,1\}$ such that

$$X_n(\omega) = \mathbb{1}_{[\frac{i_n-1}{k_n},\frac{i_n}{k_n}]}(\omega).$$

For any $\omega \in [0,1]$, we have $X_n(\omega) = 1$ for infinitely many values since there exist infinitely many (i,k) pairs such that $\frac{(i-1)}{k} \leqslant \omega \leqslant \frac{i}{k}$, and hence $\limsup_n X_n(\omega) = 1$ and hence $\lim_n X_n(\omega) \neq 0$. However, $\lim_n X_n(\omega) = 0$ in probability, since

$$\lim_n \lambda\{X_n(\omega) \neq 0\} = \lim_n \frac{1}{k_n} = 0.$$

5 Infinitely often and all but finitely many

Lemma 5.1 (infinitely often and all but finitely many). *Let $A \in \mathcal{F}^{\mathbb{N}}$ be a sequence of events.*

(a) For some subsequence $(k_n : n \in \mathbb{N})$ depending on ω, we have

$$\limsup_n A_n = \{\omega \in \Omega : \omega \in A_{k_n} \text{ for all } n \in \mathbb{N}\} = \left\{\omega \in \Omega : \sum_{n\in\mathbb{N}} \mathbb{1}_{A_n}(\omega) = \infty\right\} = \{A_n \text{ infinitely often }\}.$$

(b) For a finite $n_0(\omega) \in \mathbb{N}$ depending on ω, we have

$$\liminf_n A_n = \{\omega \in \Omega : \omega \in A_n \text{ for all } n \geqslant n_0(\omega)\} = \left\{\omega \in \Omega : \sum_{n\in\mathbb{N}} \mathbb{1}_{A_n^c}(\omega) < \infty\right\} = \{A_n \text{ for all but finitely many } n\}.$$

Proof. Let $A \in \mathcal{F}^{\mathbb{N}}$ be a sequence of events.

(a) Let $\omega \in \limsup_n A_n = \cap_{n \in \mathbb{N}} \cup_{k \geqslant n} A_k$, then $\omega \in \cup_{k \geqslant n} A_k$ for all $n \in \mathbb{N}$. Therefore, for each $n \in \mathbb{N}$, there exists $k_n \in \mathbb{N}$ such that $\omega \in A_{k_n}$, and hence

$$\sum_{j \in \mathbb{N}} \mathbb{1}_{A_j}(\omega) \geqslant \sum_{n \in \mathbb{N}} \mathbb{1}_{A_{k_n}}(\omega) = \infty.$$

Conversely, if $\sum_{j \in \mathbb{N}} \mathbb{1}_{A_j}(\omega) = \infty$, then for each $n \in \mathbb{N}$ there exists a $k_n \in \mathbb{N}$ such that $\omega \in A_{k_n}$ and hence $\omega \in \cup_{k \geqslant n} A_k$ for all $n \in \mathbb{N}$.

(b) Let $\omega \in \liminf_n A_n = \cup_{n \in \mathbb{N}} \cap_{k \geqslant n} A_k$, then there exists $n_0(\omega)$ such that $\omega \in A_n$ for all $n \geqslant n_0(\omega)$. Conversely, if $\sum_{j \in \mathbb{N}} \mathbb{1}_{A_j^c}(\omega) < \infty$, then there exists $n_0(\omega)$ such that $\omega \in A_n$ for all $n \geqslant n_0(\omega)$.

□

Theorem 5.2 (Convergence a.s. implies in probability). *If a sequence of random variables $X : \Omega \to \mathbb{R}^{\mathbb{N}}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ converges a.s. to a random variable $X_\infty : \Omega \to \mathbb{R}$, then it converges in probability to the same random variable.*

Proof. Let $\lim_n X_n = X_\infty$ a.s. and $\epsilon > 0$. We define events $A_n \triangleq \{\omega \in \Omega : |X_n(\omega) - X_\infty(\omega)| > \epsilon\}$ for each $n \in \mathbb{N}$. We will show that $\lim_n P(A_n) = 0$. To this end, let N be the exception set such that

$$N \triangleq \left\{\omega \in \Omega : \liminf_n X_n(\omega) < \limsup_n X_n(\omega) \text{ or } \limsup_n X_n(\omega) = \infty\right\}.$$

For $\omega \notin N$, there exists an $n_0(\omega)$ such that $|X_n - X_\infty| \leqslant \epsilon$ for all $n \geqslant n_0$. That is, $\omega \in A_n^c$ for all $n \geqslant n_0(\omega)$ and hence $N^c \subseteq \liminf_n A_n^c$. It follows that $1 = P(\liminf_n A_n^c)$. Since $\liminf_n A_n^c = (\limsup_n A_n)^c$, we get $0 = P(\limsup_n A_n) = \lim_n P(\cup_{k \geqslant n} A_k) \geqslant \lim_n P(A_n) \geqslant 0$. □

6 Borel-Cantelli Lemma

Proposition 6.1 (Borel-Cantelli Lemma). *Let $A \in \mathcal{F}^{\mathbb{N}}$ be a sequence of events such that $\sum_{n \in \mathbb{N}} P(A_n) < \infty$, then $P\{A_n \text{ i.o.}\} = 0$.*

Proof. We can write the probability of infinitely often occurrence of A_n, by the continuity and sub-additivity of probability as

$$P(\limsup_n A_n) = \lim_n P(\cup_{k \geqslant n} A_k) \leqslant \lim_n \sum_{k \geqslant n} P(A_k) = 0.$$

The last equality follows from the fact that $\sum_{n \in \mathbb{N}} P(A_n) < \infty$. □

Proposition 6.2 (Borel zero-one law). *Let $A \in \mathcal{F}^{\mathbb{N}}$ be a sequence of independent events, then*

$$P\{A_n \text{ i.o.}\} = \begin{cases} 0, & \text{iff } \sum_n P(A_n) < \infty, \\ 1, & \text{iff } \sum_n P(A_n) = \infty. \end{cases}$$

Proof. Let $A \in \mathcal{F}^{\mathbb{N}}$ be a sequence of independent events.

(a) From Borel-Cantelli Lemma, if $\sum_n P(A_n) < \infty$ then $P\{A_n \text{ i.o.}\} = 0$.

(b) Conversely, suppose $\sum_n P(A_n) = \infty$, then $\sum_{k \geqslant n} P(A_k) = \infty$ for all $n \in \mathbb{N}$. From the definition of lim sup and lim inf, continuity of probability, and independence of sequence of events $A \in \mathcal{F}^{\mathbb{N}}$, we get

$$P\{A_n \text{ i.o.}\} = 1 - P(\liminf_n A_n^c) = 1 - \lim_n \lim_m P(\cap_{k=n}^m A_k^c) = 1 - \lim_n \lim_m \prod_{k=n}^m (1 - P(A_k)).$$

Since $1 - x \leqslant e^{-x}$ for all $x \in \mathbb{R}$, from the above equation, the continuity of exponential function, and the hypothesis, we get

$$1 \geqslant P\{A_n \text{ i.o.}\} \geqslant 1 - \lim_n \lim_m e^{-\sum_{k=n}^m P(A_k)} = 1 - \lim_n \exp(-\sum_{k \geqslant n} P(A_k)) = 1.$$

□

Example 6.3 (Convergence in probability can imply almost sure convergence). Consider a random Bernoulli sequence $X:\Omega\to\{0,1\}^{\mathbb{N}}$ defined on the probability space $(\Omega,\mathcal{F},P)$ such that $P\{X_n=1\}=p_n$ for all $n\in\mathbb{N}$. Note that the sequence of random variables is not assumed to be independent, and definitely not identical. If $\lim_n p_n=0$, then we see that $\lim_n X_n=0$ in probability.

In addition, if $\sum_{n\in\mathbb{N}} p_n<\infty$, then $\lim_n X_n=0$ a.s. To see this, we define event $A_n\triangleq\{X_n=1\}\in\mathcal{F}$ for each $n\in\mathbb{N}$. Then, applying the Borel-Cantelli Lemma to sequence of events $A\in\mathcal{F}^{\mathbb{N}}$, we get

$$1=P((\limsup_n A_n)^c)=P(\liminf_n A_n^c).$$

That is, $\lim_n X_n=0$ for $\omega\in\liminf_n A_n^c$, implying almost sure convergence.

Theorem 6.4. *A random sequence $X:\Omega\to\mathbb{R}^{\mathbb{N}}$ converges to a random variable $X_\infty:\Omega\to\mathbb{R}$ in probability, then there exists a subsequence $(n_k:k\in\mathbb{N})\subset\mathbb{N}$ such that $(X_{n_k}:k\in\mathbb{N})$ converges almost surely to X_∞.*

Proof. Letting $n_1=1$, we define the following subsequence and event recursively for each $j\in\mathbb{N}$,

$$n_j\triangleq\inf\left\{N>n_{j-1}:P\left\{|X_r-X_\infty|>2^{-j}\right\}<2^{-j},\text{ for all } r\geqslant N\right\},\qquad A_j\triangleq\left\{\left|X_{n_{j+1}}-X_\infty\right|>2^{-j}\right\}.$$

From the construction, we have $\lim_k n_k=\infty$, and $P(A_j)<2^{-j}$ for each $j\in\mathbb{N}$. Therefore, $\sum_{k\in\mathbb{N}}P(A_k)<\infty$, and hence by the Borel-Cantelli Lemma, we have $P(\limsup_k A_k)=0$. Let $N=\limsup_k A_k$ be the exception set such that for any outcome $\omega\notin N$, for all but finitely many $j\in\mathbb{N}$

$$\left|X_{n_j}(\omega)-X_\infty(\omega)\right|\leqslant 2^{-j}.$$

That is, for all $\omega\notin N$, we have $\lim_n X_n(\omega)=X_\infty(\omega)$. □

Theorem 6.5. *A random sequence $X:\Omega\to\mathbb{R}^{\mathbb{N}}$ converges to a random variable X_∞ in probability iff each subsequence $(X_{n_k}:k\in\mathbb{N})$ contains a further subsequence $(X_{n_{k_j}}:j\in\mathbb{N})$ converges almost surely to X_∞.*

A Limits of sequences

Definition A.1. For any real valued sequence $a\in\mathbb{R}^{\mathbb{N}}$, we can define

$$\limsup_n a_n\triangleq\inf_n\sup_{k\geqslant n}a_k,\qquad\qquad \liminf_n a_n\triangleq\sup_n\inf_{k\geqslant n}a_k.$$

Remark 4. We define $e_n\triangleq\sup_{k\geqslant n}a_k$ and $f_n\triangleq\inf_{k\geqslant n}a_k$, and observe that $f_n\leqslant a_k$ for all $k\geqslant n$. That is, $f_1,\dots,f_{n-1}\leqslant a_n$ and $f_k\leqslant a_k$ for all $k\geqslant n$. It follows that $\sup_n f_n\leqslant\sup_{k\geqslant n}a_k=e_n$ for all $n\in\mathbb{N}$, and hence $\liminf_n a_n=\sup_n f_n\leqslant\inf_n e_n=\limsup_n a_n$.

Definition A.2. A sequence $a\in\mathbb{R}^{\mathbb{N}}$ is said to converge if $\limsup_n a_n=\liminf_n a_n$ and the limit is defined as $a_n\triangleq\lim_n a_n=\limsup_n a_n=\liminf_n a_n$.

Theorem A.3. *A sequence $a\in\mathbb{R}^{\mathbb{N}}$ converges to $a_\infty\in\mathbb{R}$ if for all $\epsilon>0$ there exists an integer $N\in\mathbb{N}$ such that for all $n\geqslant N$, we have $|a_n-a_\infty|<\epsilon$.*

Proof. Let $\epsilon>0$ and find the integer $N_\epsilon\in\mathbb{N}$ such that $a_n\in(a_\infty-\epsilon,a_\infty+\epsilon)$ for all $n\geqslant N_\epsilon$. It follows that $a_\infty-\epsilon\leqslant f_n\leqslant e_n\leqslant a_\infty+\epsilon$ for all $n\geqslant N_\epsilon$, and hence $a_\infty-\epsilon\leqslant\liminf_n a_n\leqslant\limsup_n a_n\leqslant a_\infty+\epsilon$. Since ϵ was arbitrary, it follows that $\lim_n a_n=a_\infty$. □

Proposition A.4. *For any sequence $a \in \mathbb{R}_+^{\mathbb{N}}$, the following statements are true.*

(i) If $\sum_{n\in\mathbb{N}} a_n < \infty$ then $\lim_{n\to\infty} \sum_{k\geqslant n} a_k = 0$.

(ii) If $\sum_{n\in\mathbb{N}} a_n = \infty$ then $\sum_{k\geqslant n} a_k = \infty$ for all $k \in \mathbb{N}$.

Proof. We observe that $(\sum_{k<n} a_k : n \in \mathbb{N})$ is a non-decreasing sequence, and hence $\lim_{n\to\infty} \sum_{k<n} a_k = \sup_n \sum_{k<n} a_k = \sum_{n\in\mathbb{N}} a_n$.

(i) It follows that $\sum_{k\geqslant n} a_k = \sum_{n\in\mathbb{N}} a_n - \sum_{k<n} a_k$ is a non-increasing sequence with limit 0.

(ii) We can write $\sum_{n\in\mathbb{N}} a_n = \sum_{k<n} a_k + \sum_{k\geqslant n} a_k$. Since the first term is finite for all $n \in \mathbb{N}$, it follows that the second term must be infinite for all $n \in \mathbb{N}$.

□