Lecture-14: Almost sure convergence of random variables

1 Almost sure convergence

Consider a probability space $(\Omega, \mathcal{F}, P)$. Recall that a random variable X is an $\mathcal{F}$-measurable function on the sample space Ω such that $X^{-1}(-\infty, x] \in \mathcal{F}$ for all $x \in \mathbb{R}$. A sequence of random variables $(X_n(\omega) : n \in \mathbb{N})$ is hence a sequence of $\mathcal{F}$-measurable functions. There are many possible definitions for convergence of a sequence of random variables. One idea is to require $X_n(\omega)$ to converge for each fixed ω. However, at least intuitively, what happens on an event of probability zero is not important.

Definition 1.1. A statement holds **almost surely** (abbreviated a.s.) if there exists an event called the **exception set** $N \in \mathcal{F}$ with $P(N) = 0$ such that the statement holds if $\omega \notin N$.

Example 1.2 (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 there exists an exception set

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

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

Example 1.3 (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 there exists an exception set $N = \{\omega \in \Omega : X(\omega) > Y(\omega)\} \in \mathcal{F}$ and $P(N) = 0$.

Definition 1.4. If $(X_n : n \in \mathbb{N})$ is a sequence of random variables, then $\lim_n X_n$ exists a.s. means there exists an exception event $N \in \mathcal{F}$, such that $P(N) = 0$ and if $\omega \notin N$, then $\lim_n X_n(\omega)$ exists. That is,

$$N^c = \left\{\omega \in \Omega : \limsup_n X_n(\omega) = \liminf_n X_n(\omega)\right\}.$$

Let X be the point-wise limit of the sequence of random variables $(X_n(\omega) : n \in \mathbb{N})$ on the set N^c, then we say that the sequence $(X_n(\omega) : n \in \mathbb{N})$ **converges almost surely** to X, and denote it as

$$\lim_n X_n = X \text{ a.s.}$$

Example 1.5 (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 $0 \leqslant a \leqslant b \leqslant 1$. We define the scaled indicator random variable $X_n : \Omega \to \{0,1\}$ such that

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

Let $N = \{0\}$, then for any $\omega \notin N$, there exists $m = \lceil \frac{1}{\omega} \rceil \in \mathbb{N}$, such that for all $n \geqslant m$, 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 $n \in \mathbb{N}$.

2 Convergence in probability

Definition 2.1. A sequence $(X_n : n \in \mathbb{N})$ of random variables **converges in probability** to a random variable X, if for any $\epsilon > 0$

$$\lim_n P\{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon\} = 0.$$

Remark 1. For a sequence $(X_n : n \in \mathbb{N})$, almost sure convergence of means that for almost all outcomes ω, the difference $X_n(\omega) - X(\omega)$ gets small and stays small. Convergence in probability is weaker and merely requires that the probability of the difference $X_n(\omega) - X(\omega)$ being non-trivial becomes small.

Example 2.2 (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, such that $n = S_{k-1} + i$ for $i \in [k]$. Therefore, for each $n \in \mathbb{N}$, we define indicator random variable $X_n : \Omega \to \{0,1\}$ such that

$$X_n(\omega) = \mathbb{1}_{\left[\frac{i-1}{k}, \frac{i}{k}\right]}(\omega).$$

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

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

3 Borel-Cantelli Lemma

Lemma 3.1 (infinitely often and almost all). *Let $(A_n \in \mathcal{F} : n \in \mathbb{N})$ be a sequence of events.*

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

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

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

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

Proof. Let $(A_n \in \mathcal{F} : n \in \mathbb{N})$ be a sequence of events.

(a) Let $\omega \in \limsup_n A_n = \cap_{n \in \mathbb{N}} \sup_{k \geqslant n} A_k$, then $\omega \in \sup_{k \geqslant n} A_k$ 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 $\omega \in \liminf_n A_n$, 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 3.2 (Convergences a.s. implies in probability). *If a sequence of random variables $(X_n(\omega) : n \in \mathbb{N})$ defined on a probability space $(\Omega, \mathcal{F}, P)$ converges a.s. to a random variable X, then it converges in probability to the same random variable.*

Proof. Since $\lim_n X_n = X$ a.s., let N be the exception set. Let $\epsilon > 0$ and $\omega \notin N$, then there exists an $n_0(\omega)$ such that $|X_n - X| \leqslant \epsilon$ for all $n \geqslant n_0$. Defining event $A_n = \{\omega \in \Omega : |X_n(\omega) - X(\omega)| > \epsilon\}$, we have

$$1 = P(A_n^c \text{ for all but finitely many } n) = P(\liminf_n A_n^c).$$

Since $\liminf_n A_n^c = (\limsup_n A_n)^c$, we get

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

□

Proposition 3.3 (Borel-Cantelli Lemma). *Let $(A_n \in \mathcal{F} : n \in \mathbb{N})$ 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 3.4 (Borel zero-one law). *If $(A_n \in \mathcal{F} : n \in \mathbb{N})$ is 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 $(A_n \in \mathcal{F} : n \in \mathbb{N})$ 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 events $(A_k \in \mathcal{F} : k \in \mathbb{N})$ 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 3.5 (Convergence in probability can imply almost sure convergence). Consider a sequence of Bernoulli random variables $(X_n \in \{0,1\} : n \in \mathbb{N})$ 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 observe from application of the Borel-Cantelli Lemma to events $(A_n = \{X_n = 1\} : n \in \mathbb{N})$ that

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

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

Theorem 3.6. *A random sequence $(X_n : n \in \mathbb{N})$ converges to a random variable X 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 $j \in \mathbb{N}$,

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

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 $j \in \mathbb{N}$

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

That is, for all $\omega \notin N$, we have $\lim_n X_n(\omega) = X(\omega)$. ☐

Theorem 3.7. *A random sequence $(X_n : n \in \mathbb{N})$ 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.*